

UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR



TRABAJO FIN DE MÁSTER

DESARROLLO DE UN MODELO PREDICTIVO DEL ÉXITO DE INDUCCIÓN DEL PARTO

Doble Máster en Ingeniería Informática e Investigación e Innovación en Tecnologías de la Información y las Comunicaciones

Cristina Pruenza García-Hinojosa

Tutores: Julia Díaz García y Ana González Marcos

Departamento de Ingeniería Informática

Escuela Politécnica Superior

Universidad Autónoma de Madrid

Junio 2016

DESARROLLO DE UN MODELO PREDICTIVO DEL ÉXITO DE INDUCCIÓN DEL PARTO

Cristina Pruenza García-Hinojosa
Tutores: Julia Díaz García y Ana González Marcos

Departamento de Ingeniería Informática
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Junio 2016

Resumen

En la actualidad, la inducción del parto es una de las intervenciones mejor estudiadas dentro del campo de la Obstetricia, pero todavía no se han identificado buenos factores predictivos del éxito de este procedimiento. Este hecho obliga al médico experto a tomar la decisión de inducir en base a ciertas características de la madre y/o el feto, según su criterio, pero que en el caso de ser una decisión errónea puede desencadenar en la realización de una cesárea e, incluso, otras complicaciones indeseadas. Es por ello que uno de los retos en Obstetricia consiste en la mejora de la capacidad predictiva de éxito de una inducción. En este campo resulta de vital importancia la adquisición de datos fiables y representativos, pues a pesar de que existe gran cantidad de información procedente de diversas fuentes, estructurarla e inferir decisiones no siempre es sencillo. En este punto es donde el Big Data y el aprendizaje automático están adquiriendo cada vez más importancia.

En este proyecto se ha propuesto un sistema robusto para explotar y modelar una fuente de información clínica, extrayendo valor de ella. Con el fin de satisfacer el principal reto del proyecto, se han diseñado diferentes modelos predictivos del éxito de un procedimiento de inducción del trabajo de parto. En primer lugar, se ha modelado el conocimiento experto mediante reglas de decisión que conforman la base del sistema de referencia. En segundo lugar, se han utilizado diferentes algoritmos de aprendizaje automático y modelos de paciente para inferir esas mismas decisiones, a fin de obtener un sistema con mayor riqueza. En esta misma línea, se estudia un método general para explicar las predicciones emitidas por los clasificadores, dotando así de mayor transparencia a los algoritmos. Los resultados obtenidos tanto en el sistema de recuperación de información como en las predicciones y explicaciones de la clasificación son considerados satisfactorios y muestran la efectividad y robustez de la herramienta.

Al finalizar este proyecto se ha presentado un sistema de utilidad para el experto en la ayuda a la toma de decisiones, con capacidad de proporcionar una explicación a las mismas y con ello fomentar su uso en el ámbito médico. Aparte de esto, se ha diseñado un sistema de validación y visualización de los datos clínicos estructurados, el cual permite inferir resultados directamente del sistema de referencia.

Palabras Clave

Aprendizaje automático, Big Data, Historia Clínica Electrónica, inducción del parto, modelo predictivo, modelos de paciente, motor de inferencia, reglas de decisión.

Abstract

Nowadays, induction of labour is one of the best studied procedures within the area of Obstetrics. However, no positive prediction values of the success of this procedure have been identified yet. This fact forces skilled specialists to make the decision to induce labour, based on the characteristics of the mother and/or foetus and on their own criteria, but if this decision is incorrect, it may lead to a caesarean section or even some unnecessary medical complications. This fact leads to one of the main challenges in Obstetrics consisting of improving the prediction of successful induction of labour. In this area, the importance of obtaining reliable and representative data is vital. Although it is possible to find a large amount of information from different sources, it is not always easy to structure it and infer knowledge from it. This is the reason why Big Data and machine learning techniques are becoming more and more important in this area.

For this project we proposed a robust method for exploring and modeling a source of clinical information to obtain all the possible knowledge. In order to address the main challenge of this project, different predictive models for induction success of labour were designed.

Firstly, skilled knowledge was adapted using decision-making rules, which are the basis for the reference system.

Secondly, different machine learning algorithms and patient models were used in order to obtain the same decisions, and to create a more enriched system.

Along this same line, a general method is being studied to explain the predictions emitted by the classifiers, thus providing the algorithm with more transparency. The positive results obtained from both the information recovery system and the predictions and explanations of the classification, show the effectiveness and strength of this tool.

Finally, a utility system for the skilled specialist was presented in order to help making the decision. This system has the capacity of proposing an explanation and therefore encouraging its use in the medical field. Furthermore, a validation and visualization of the structured clinic data system has been designed, allowing deductions of results to be made directly from the reference system.

Key words

Machine learning, Big Data, Electronic Medical Records, induction of labor, predictive model, patient models, inference engine, decision-making rules.

Agradecimientos

Me gustaría que estas líneas sirvieran para expresar mi agradecimiento a todas aquellas personas que con su ayuda han colaborado en la realización del presente trabajo.

En primer lugar quería agradecer a Julia Díaz García y a Ana González Marcos, el apoyo y la confianza que me han transmitido para llevar a cabo este proyecto. Y al Instituto de Ingeniería del Conocimiento así como al Hospital Universitario de Fuenlabrada por permitir que esto fuese posible.

Agradecer a Luis Lechuga, a la doctora María Teulón, a Carmen Yelo y a Marta Dávila su activa colaboración en el proyecto, por facilitar el conjunto de datos INDUCESS y dedicar su valioso tiempo en todas las revisiones y reuniones mantenidas.

Quiero dar las gracias a toda la gente que ha estado dispuesta a ayudarme, en especial a mis compañeros del IIC, Álvaro, Edoardo, Sandra, Dani, Álvaro y Jose, y a los que ya no están, Ángela y Javi, ya que su labor ha sido imprescindible para llegar hasta aquí.

A mis amigos, Julia, Rocío y Pencho porque han estado siempre demostrándome su apoyo y son muchos los momentos vividos juntos.

Agradecer también a toda mi familia, a mis padres, a mi abuela, a mi hermano, a Mamen y al peque de la casa, Hugo, por todo lo que hacen para animarme, apoyarme y darme fuerzas cada día, sin ellos nada de esto hubiese sido posible.

Pero si hay alguien a quien debo dar las gracias es, sin duda, a Andrés porque solamente él sabe el trabajo que ha supuesto este proyecto, su enorme paciencia y ayuda han sido esenciales para que hoy esté donde estoy. Es el mejor regalo que me llevo de mi paso por esta facultad.

A todos los demás, que de una forma u otra han sido parte de este trabajo, gracias.

*Cristina Pruenza García-Hinojosa
Junio 2016*

Índice general

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.3. Estructura del documento	3
2. Estado del arte	5
2.1. Inducción del trabajo de parto	5
2.2. Sistemas de información clínica	7
2.3. Sistemas basados en conocimiento	9
2.3.1. Sistemas expertos basados en reglas	10
2.3.2. Ontologías	12
2.4. Explicación en clasificación	13
3. Sistema de recuperación de información	19
3.1. Conjunto de datos	19
3.1.1. Preproceso de datos	20
3.1.2. Diseño del modelo de datos	21
3.2. Diseño del sistema IR	22
3.3. Sistema de validación	25
3.3.1. Diseño de la interfaz web	26
4. Sistema basado en reglas	29
4.1. Análisis y diseño de reglas	29
4.1.1. Tratamiento de prioridades	30
4.1.2. Verificación de reglas	32
4.2. Diseño del sistema experto	33
5. Big Data en medicina	37
5.1. Modelos utilizados	38
5.1.1. Naive Bayes	38
5.1.2. Árboles de decisión	38
5.1.3. Redes neuronales	39

5.1.4. Máquinas de vectores soporte	40
5.1.5. Random Forest	40
5.1.6. Vecinos más cercanos	41
5.1.7. Maximización de la esperanza	41
5.2. Selección de atributos	42
5.2.1. Medidas de evaluación	42
5.2.2. Algoritmos de selección	43
6. Evaluación y análisis de resultados	45
6.1. Preparación de datos	46
6.1.1. Selección de atributos	47
6.2. Métricas de evaluación	49
6.3. Resultados de validación del sistema IR	51
6.4. Resultados de inferencia	53
6.4.1. Resultados del sistema experto	53
6.4.2. Resultados de los modelos de aprendizaje automático	54
6.4.3. Evaluación comparativa	58
6.5. Explicación en clasificación	60
7. Conclusiones y trabajo futuro	71
7.1. Conclusiones	71
7.2. Trabajo futuro	73
Bibliografía	77
Glosario	79
Acrónimos	81
A. Descripción de atributos	83
B. Pantallas página web	89
C. Reglas de decisión	95
D. Matrices de confusión	99

Índice de figuras

2.1. Arquitectura de un sistema experto.	9
3.1. Grafo de atributos de un paciente.	21
3.2. Proceso seguido en el sistema IR.	22
3.3. Pantalla inicial del sistema de validación.	26
3.4. Pantalla para validación del sistema.	27
4.1. Estructura general del sistema experto diseñado.	33
4.2. Mecanismo de inferencia del sistema experto.	35
6.1. Comparativa del error de clasificación entre los Modelos 2 y cada submodelo del Modelo 3 para los distintos algoritmos utilizados.	59
6.2. Representación bidimensional del conjunto de datos INDUCESS.	61
6.3. Contribuciones globales de los atributos a la predicción de cada clase para una red neuronal.	62
6.4. Histograma de <i>motivo_induccion</i> en el conjunto de datos INDUCESS.	63
6.5. Contribuciones de diferentes valores de un atributo a la predicción de las clases para una red neuronal.	64
6.6. Contribuciones globales de los atributos a la predicción de cada clase para una SVM.	66
6.7. Contribuciones de diferentes valores de un atributo a la predicción de la clase <i>Inducción</i>	67
6.8. Contribuciones globales de los atributos a la predicción de cada clase para un árbol.	68
6.9. Contribuciones de diferentes valores de un atributo a la predicción de las clases para un árbol.	69
A.1. Histogramas de variables predictoras y clase del conjunto de datos INDUCESS (Parte 1).	86
A.2. Histogramas de variables predictoras y clase del conjunto de datos INDUCESS (Parte 2).	87

A.3. Histogramas de variables predictoras y clase del conjunto de datos INDUCESS (Parte 3).	87
B.1. Pantalla inicial completa del sistema de validación.	90
B.2. Pantalla de acceso al sistema de validación.	91
B.3. Pantalla de resultado de INDUCESS.	92
B.4. Pantalla de estadísticas del sistema de validación.	93
B.5. Pantalla de descripción del equipo del proyecto.	94

Índice de tablas

3.1. Atributos de la tabla <i>pacientes</i> de la base de datos.	25
4.1. Distribución por pesos de los atributos predictores del modelo.	31
6.1. Atributos que componen el Modelo 1 (Falsos predictores).	46
6.2. Atributos que componen el Modelo 2 (Atributos predictores).	47
6.3. Atributos que componen el Modelo 3 (Atributos predictores tras selección). . . .	48
6.4. Modelos de paciente utilizados para técnicas de aprendizaje automático.	49
6.5. Esquema de los C problemas de clasificación binaria a realizar.	50
6.6. Representación de la matriz de confusión para cada problema de clasificación binaria.	51
6.7. Porcentajes de error en el sistema de recuperación de información.	52
6.8. Errores más comunes en las variables extraídas.	52
6.9. Resultados del sistema experto (Modelo de referencia).	53
6.10. Resultados de los algoritmos de clasificación del Modelo 1 (Falsos predictores). .	55
6.11. Resultados de los algoritmos de clasificación del Modelo 2 (Atributos predictores). 55	
6.12. Resultados de los algoritmos de clasificación del Modelo 3 (Atributos predictores tras selección).	56
6.13. Resultados de los algoritmos de <i>clustering</i> del Modelo 1 (Falsos predictores) y Modelo 2 (Atributos predictores).	57
6.14. Resultados de los algoritmos de <i>clustering</i> del Modelo 3 (Atributos predictores tras selección).	57
D.1. Matrices de confusión del Modelo 1 (Falsos predictores).	99
D.2. Matrices de confusión del Modelo 2 (Atributos predictores).	100
D.3. Matrices de confusión del Modelo 3 - Atributos predictores tras ReliefF.	100
D.4. Matrices de confusión del Modelo 3 - Atributos predictores tras mRMR.	101
D.5. Matrices de confusión del Modelo 3 - Atributos predictores tras GainRatioAttri- buteEval.	101
D.6. Matrices de confusión del Modelo 3 - Atributos predictores tras InfoGainAttri- buteEval.	102

D.7. Matrices de confusión del Modelo 3 - Atributos predictores tras CorrelationAttributeEval.	102
--	-----

1

Introducción

1.1. Motivación

El embarazo es uno de los períodos más críticos en la vida de una mujer. A lo largo de esta etapa pueden aparecer ciertos problemas y conlleven la necesidad de provocar el parto con el fin de extraer el feto con éxito y proteger a la madre o a ambos. Es en estos casos donde la inducción del parto desempeña un papel tan importante, ya que las circunstancias obligan al obstetra a actuar siendo imposible seguir el curso fisiológico del mismo. El objetivo de la inducción del trabajo de parto es el de estimular las contracciones uterinas antes de su inicio espontáneo para tener como resultado un parto vaginal con éxito.

En las últimas dos décadas la tasa de inducción del parto se ha duplicado convirtiéndose en un procedimiento bastante común, empleado en más de un 20 % de las gestaciones [1] [2]. A pesar de que la inducción sea una de las intervenciones mejor estudiadas dentro del campo de la Obstetricia y existan numerosos estudios publicados acerca de ella, todavía no se han identificado buenos factores predictivos del éxito de este procedimiento. Este hecho provoca que hasta el momento no existan herramientas que apoyen al médico experto en su criterio a la hora de tomar la decisión de realizar una inducción, viéndose obligado a decidir en base a ciertas características de la madre y el feto, pero que en el caso de ser una decisión no acertada podría suponer un riesgo para ambos.

En la actualidad, existen ciertas variables maternas y fetales ligadas al éxito del procedimiento de inducción del parto pero, sin embargo, su valor predictivo es inferior a lo deseado. Esto puede llevar al experto a tomar una decisión inadecuada [2] [3] y, en muchos casos, desencadenar en la realización de una cesárea e, incluso, ocasionar otras complicaciones indeseadas. Bajo

esta situación resulta interesante conocer a priori la posibilidad de éxito de la inducción para desestimar aquellas inducciones con alta probabilidad de fracaso, evitando así procedimientos innecesarios o posponiendo la finalización del embarazo si fuese posible [4].

Por ello, mejorar la capacidad predictiva de éxito de una inducción del parto constituye uno de los retos actuales de la Obstetricia contribuyendo, a su vez, a mejorar los resultados en salud, los costes tanto en medicación, hospitalización o de personal y la calidad percibida del procedimiento.

Por otra parte, el modelo sanitario y en general el sector de la salud, es hoy en día uno de los sectores donde el Big Data está teniendo un mayor impacto y donde sus aplicaciones están creciendo de un modo exponencial, tanto para el área médica como también para las áreas de análisis de datos. Los médicos mueven grandes volúmenes de datos clínicos procedentes de distintas fuentes como, por ejemplo, sus historias clínicas electrónicas, sus análisis clínicos o los *wearables* que registran datos sobre constantes vitales de salud, los cuales pueden resultar difíciles de explotar eficaz y eficientemente. Es en este punto donde el aprendizaje automático y las técnicas de Big Data adquieren cada vez más importancia para extraer, representar y explotar el conocimiento real de estos datos, estructurados y no estructurados, con el fin de ayudar al experto en su práctica clínica.

En este marco descrito, se encuadran la motivación y los objetivos de este proyecto, desglosados a continuación.

1.2. Objetivos

El principal objetivo de este proyecto es desarrollar un modelo predictivo del éxito de un procedimiento de inducción del trabajo de parto y construir una herramienta de ayuda a la toma de decisiones basada en dos enfoques. Por un lado, un sistema para inferir decisiones que se apoya en reglas de decisión provenientes de un panel de expertos en Obstetricia y, por otro, modelos de inferencia basados en algoritmos y técnicas de aprendizaje automático y Big Data, a la vez que se da continuidad a una vía de investigación consistente en la posible explicación de la clasificación obtenida [5].

Más concretamente, para alcanzar este objetivo y satisfacer las motivaciones de este proyecto se han establecido los siguientes objetivos parciales:

1. Diseñar e implementar técnicas para la recuperación de información así como una interfaz que permita visualizar de manera estructurada los datos extraídos. De esta forma se optimiza la gestión clínica y, a su vez, sirve de herramienta para la validación de datos e inferir el éxito de la inducción al parto mediante el primer enfoque descrito.
2. Modelar y acreditar la validez y confiabilidad de las reglas con las que el clínico toma la

decisión de inducir el proceso en función de unas variables y su relevancia o peso. Estas reglas, variables y peso serán definidas como el modelo de referencia por un panel de expertos en Obstetricia.

3. Aplicar técnicas de aprendizaje automático para valorar la relevancia de las variables y obtener modelos de inferencia basados en la información recopilada de los distintos pacientes. Los diferentes modelos permitirán medir la efectividad del modelo de referencia y obtener una nueva clasificación.
4. Analizar detalladamente algunas de las técnicas anteriores que facilitan su uso por parte de los especialistas médicos y con ello dotar de mayor transparencia a los modelos considerados, especialmente aquellos conocidos tradicionalmente como de *caja negra*.
5. Diseñar un módulo de evaluación o test, que permita inferir la inducción de un nuevo paciente según los distintos modelos, y evaluar posteriormente la utilidad de cada uno de ellos.

Cumpliendo todos estos objetivos se espera estar más cerca de satisfacer la motivación principal del proyecto. Este trabajo apunta a contribuir, de forma inmediata o indirecta, en tres direcciones de avance: por un lado, explotar una fuente de información clínica y extraer valor de ella; por otro, desarrollar una herramienta para la inferencia de la inducción al parto con éxito; y finalmente llevar a cabo un exhaustivo estudio sobre la utilidad de distintos modelos de aprendizaje automático que se aplican exitosamente al dominio médico así como ofrecer, de alguna forma comprensible para el especialista en salud, el conocimiento adquirido sobre el problema.

1.3. Estructura del documento

La memoria de este trabajo consta de los siguientes apartados:

- **Capítulo 1.** Introducción: motivación y objetivos del proyecto.
- **Capítulo 2.** Estado del arte: estudio del procedimiento de inducción del trabajo de parto y algunas variables predictoras, análisis de la Historia Clínica Digital y su aplicación como fuente de información, descripción de los sistemas basados en conocimiento y de técnicas para la explicación de la clasificación.
- **Capítulo 3.** Sistema de recuperación de información: análisis de los atributos disponibles, sistema de extracción de la información utilizada en la inferencia y sistema de validación.
- **Capítulo 4.** Sistema basado en reglas: análisis de prioridades, verificación de las reglas que componen el sistema experto y diseño del mismo.

- **Capítulo 5.** Big Data en medicina: análisis de modelos de aprendizaje automático aplicados en biomedicina y métodos de selección de atributos.
- **Capítulo 6.** Evaluación y análisis de resultados: preparación de los datos para los experimentos, métricas de validación y evaluación, y resultados. Explicación de la clasificación obtenida del conjunto de datos real sobre embarazos.
- **Capítulo 7.** Conclusiones y trabajo futuro.
- **Referencias, glosario, acrónimos y anexos.**

2

Estado del arte

En esta sección se van a introducir y a contextualizar una serie de conceptos que serán recurrentes a lo largo de este proyecto. En concreto, se describirá el procedimiento de inducción del trabajo de parto introduciendo brevemente algunos conceptos fundamentales dentro de este ámbito y se dará una visión genérica de la Historia Clínica Digital como fuente de información. Por otra parte, se hará una breve reseña que se extenderá en capítulos posteriores, sobre los sistemas expertos y las herramientas necesarias para su desarrollo, se estudiarán ciertas aproximaciones propuestas para la explicación de la clasificación basada en los valores de los atributos y así hacer más transparentes los modelos utilizados.

2.1. Inducción del trabajo de parto

El parto es el conjunto de fenómenos fisiológicos y mecánicos, que determina la expulsión de un producto viable de los genitales maternos y sus anexos. Una mujer inicia el parto con la aparición de contracciones uterinas regulares, que aumentan en intensidad y frecuencia, acompañadas de la intervención de mecanismos fisiológicos que favorecen el estado del cuello uterino.

Debido a que el inicio del trabajo de parto varía entre una mujer y otra, si éste progresa demasiado lento, o si no se ha desencadenado naturalmente por sí mismo, a veces es necesaria su estimulación mediante métodos que se caracterizan por la iniciación artificial de la actividad uterina [3]. Dicha estimulación es lo que se conoce como inducción del trabajo de parto o, comúnmente, inducción del parto, la cual se realiza con el objetivo de iniciar o acelerar el trabajo de parto para tener como resultado un parto vaginal. Se debe diferenciar de la conducción, que significa mejorar la labor de parto cuando éste ya se ha iniciado espontáneamente, es decir, ya

existen contracciones y dilatación a nivel del cuello uterino [6].

Aunque todavía no se han identificado evidentes factores predictivos del éxito del procedimiento de inducción del trabajo de parto, existen ciertas situaciones obstétricas en las cuales este procedimiento se considera más conveniente para el desenlace favorable materno-fetal y, éstas serán la base para llevar a cabo el desarrollo del sistema basado en reglas.

En los trabajos [7] y [8] se presentan variables habituales relacionadas con situaciones en las que se decide inducir el trabajo de parto y, por tanto, se consideran predictoras de la inducción. Entre otras, complicaciones durante el embarazo (diabetes, preeclampsia o eclampsia), crecimiento fetal deficiente o menor cantidad de líquido amniótico. Sin embargo, existen otras muchas que aparecen durante o después del parto, las cuales no aportan ninguna información en el proceso de la toma de decisiones y se podrían eliminar del modelo de inferencia. Estas variables son denominadas de resultado.

Puesto que la decisión de inducir proviene únicamente del conocimiento clínico en base a protocolos, guías y experiencia previas, las reglas de decisión que conforman el modelo de referencia, se definen en función de esas variables asociadas a la evolución clínica. Por ello, resulta de interés valorar la efectividad de las mismas y buscar otros modelos de inferencia que determinen las variables relevantes predictoras.

En cuanto a los métodos existentes para la inducción del trabajo de parto pueden ser de dos tipos: mecánicos o farmacológicos. Pero en la práctica obstétrica, el método preferido, con frecuencia, depende de la experiencia del experto y de la disposición y los requisitos del método. Aunque hoy en día los más empleados son los métodos farmacológicos y, por tanto, los tratados en este proyecto, destacando la utilización de prostaglandinas como el más frecuente [9]:

- **Relaxina.** Esta hormona se usa de forma experimental. Inicialmente, se empleaba relaxina de cerdo purificada, pero actualmente se emplea la forma recombinante humana.
- **Oxitocina.** Es una hormona similar a la segregada por la hipófisis posterior. Se trata de una de las formas más frecuentes de inducción del parto.
- **Prostaglandinas.** Es el método farmacológico más utilizado. Se trata de sustancias derivadas del ácido araquidónico. Existen dos tipos fundamentales de prostaglandinas:
 - (a) PGE_2 (*Dinoprostona*). Son las únicas prostaglandinas aprobadas para su uso en la maduración cervical. Entre las dos maneras de presentación que existen se encuentra el *Dispositivo vaginal de liberación controlada (Propress)*.
 - (b) PGE_1 (*Misoprostol*). Se comercializa para la prevención y tratamiento de la úlcera péptica. En la actualidad, se está desarrollando una amplia investigación para su uso en la maduración cervical y el tratamiento de la hemorragia posparto.

En general, aunque la inducción es segura, tiene algunos riesgos potenciales que varían de

acuerdo con el método utilizado y con la situación del paciente. Destaca el aumento del riesgo del parto por cesárea y del parto instrumental [1], ya sea por ventosas, fórceps o espátulas. El procedimiento de inducción suele realizarse, en un entorno hospitalario vigilado y suele ser largo (entre 12 y 24 horas) y, cuando no es exitoso, puede desencadenar en intervenciones quirúrgicas de manera inmediata. Esto es una fuente importante de quejas y reclamaciones que conviene mejorar.

Asimismo la cesárea presenta menores complicaciones cuando se realiza de forma programada que cuando se realiza con carácter urgente o emergente durante el parto. Por ello, el modelo predictivo a desarrollar podría evitar procesos con alta probabilidad de fracaso, disminuir sus complicaciones, mejorando resultados en salud y la calidad de la atención percibida de los pacientes.

2.2. Sistemas de información clínica

Los sistemas de información clínica están teniendo un gran avance tecnológico en los últimos años, permitiendo la evolución de la información sanitaria tradicional en papel hacia sistemas electrónicos. Éstos han de ser capaces de integrar todos los aspectos relevantes sobre el estado de salud de los pacientes, además de ser accesibles universalmente y garantizar su uso ético, la confidencialidad de la información y la privacidad de las personas [10]. Por tanto, todo sistema de información ha de presentar controles que protejan la privacidad de la información clínica.

A pesar de este avance, la aplicación de las Tecnologías de la Información y Comunicación (TIC) al sector sanitario está más atrasada que en otros campos, como el del comercio o la banca. Esto es debido a que el escenario es demasiado grande para que el Sistema Nacional de Salud pueda acometerla simultáneamente en su totalidad, y a que las TIC han crecido más rápido de lo que los recursos económicos del sistema sanitario han podido asumir [11].

Pero además para satisfacer los requerimientos necesarios para el manejo de tal volumen de información sanitaria, se necesitan numerosas infraestructuras. Se podría decir que se trata de un sistema de gestión del conocimiento a gran escala, por lo que las dificultades que plantea son múltiples [12].

Tal y como se ha comentado anteriormente, la situación está cambiando y la implantación de la Historia Clínica Electrónica es una de las prioridades de los servicios de salud en los últimos años [13]. La Historia Clínica Electrónica se define en [14] como el conjunto de documentos que contienen datos, valoraciones e información de cualquier índole sobre la situación y la evolución clínica de un paciente durante el proceso asistencial. Está constituida por el conjunto de documentos, tanto escritos como gráficos, que hacen referencia a los episodios de salud y enfermedad de una persona, y a la actividad sanitaria que se genera con motivo de esos episodios.

Teniendo en cuenta la complejidad del proceso de automatización de esta información, las TIC incluyen numerosas ventajas frente al sistema tradicional, incorporando nuevas formas de obtener información de forma rápida y eficaz. Además solucionan algunos problemas existentes como son la falta de uniformidad entre los modelos utilizados en diferentes centros sanitarios, la ilegibilidad de la información escrita a mano, la accesibilidad, la necesidad de espacio para su almacenamiento, las dudosas garantías de confidencialidad y, además, la posibilidad de que su contenido pueda alterarse, perderse o deteriorarse se reduce.

La información de la Historia Clínica de los pacientes se registra y almacena principalmente en texto libre en campos de información poco estructurada. Por lo que ésta podría ser contradictoria o ambigua, siendo necesario desarrollar técnicas y algoritmos que transformen los datos en conocimiento estructurado y normalizado. Y con esto, dar soporte a los procesos de analítica avanzada y a la toma de decisiones clínicas. Este proceso será explicado en mayor detalle en la Sección 3.

Entre los diferentes sistemas de información asistencial, se encuentra Selene [15], que es una solución clínica integral que gestiona el proceso asistencial completo. Selene almacena la información generada en forma de objetos clínicos, independientemente de dónde o quién la haya generado. Se trata de un cambio conceptual de «orientación a proceso», que permite la definición de un producto único integrando funcionalidad clínica y administrativa para completar los diferentes trabajos de cada organización. Asimismo, el concepto de proceso asistencial permite asentar las bases para la orientación hacia un sistema potenciado con sistemas de ayuda a la toma de decisiones.

Los objetos clínicos, citados anteriormente, son visibles y explotables con independencia del proceso en el que se han generado. Son la unidad de información que se vincula con el proceso. Hay cuatro tipos generales: *informes*, *notas*, *formularios* y *peticiones*. A través de ellos el personal clínico realiza su registro de actuación sobre el paciente. Es decir, son los componentes mínimos de la Historia Clínica Electrónica.

Siempre teniendo en cuenta la premisa de normalización, la introducción de información en la Historia Clínica del paciente se realiza por tres vías básicas que son: *Notas de progreso*, *Hojas de Evolución*, *Informes y Peticiones*.

- **Notas de progreso:** objetos de información en texto libre, categorizados en varios tipos definibles por el usuario. Un ejemplo es la *Nota de Observaciones de Enfermería*, *Evolución del paciente* o *Anamnesis*, entre otros.
- **Hojas de evolución:** formularios de captura de datos estructurados que alimentan los objetos clínicos. Aquí se incluyen los registros de *Constantes vitales*, *balance hídrico* o *medidas antropométricas*.
- **Informes y Peticiones:** los informes son documentos contruidos con un procesador de

texto a partir de plantillas previamente definidas. Estos documentos pueden ser generados de manera automática a partir de formularios o de datos de paciente ya recogidos en el sistema. En cuanto a las peticiones, se pueden realizar a departamentos diagnósticos internos o externos, tratamientos o solicitudes a otros usuarios del sistema. Además de visualizar el estado de la petición.

2.3. Sistemas basados en conocimiento

Los sistemas expertos o sistemas basados en conocimiento son programas diseñados para actuar como un especialista humano en un dominio particular o área de conocimiento. En este sentido, pueden considerarse como intermediarios entre el experto humano, que transmite su conocimiento al sistema, y el usuario que lo utiliza para resolver un problema con la eficacia del especialista.

Los sistemas basados en conocimiento se pueden considerar como el primer producto verdaderamente operacional de la inteligencia artificial. En primer lugar, son generalmente programas útiles y prácticos que colman una necesidad concreta en cualquier área del conocimiento humano y, en segundo lugar, factibles y viables económicamente para las organizaciones. El objetivo de éstos es ayudar en la toma de decisiones pero nunca sustituir al experto, pues es él quien debe tener la última palabra.

Un sistema experto está conformado por *especialistas humanos* e *ingenieros del conocimiento*. Además consta de una *base de conocimiento*, que contiene conocimiento modelado extraído del diálogo con un experto, de una *base de hechos* o *memoria de trabajo*, que son los hechos sobre un problema que se ha descubierto durante el análisis y el *motor de inferencia*, que modela el proceso de razonamiento humano. Junto a esto, también existen *módulos de justificación*, que explican el razonamiento utilizado por el sistema para llegar a una determinada conclusión, y la *interfaz de usuario*, que es la interacción entre el sistema y el usuario, y se realiza mediante el lenguaje natural. El motor de inferencia es el encargado de determinar las reglas que se aplican en cada momento y de ejecutarlas. En la Figura 2.1 se muestra un esquema de esta arquitectura descrita.

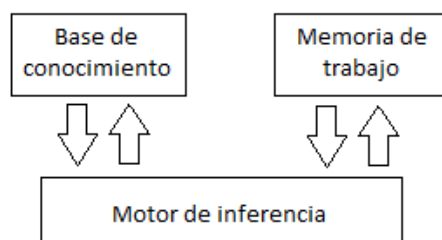


Figura 2.1: Arquitectura de un sistema experto.

Principalmente existen tres tipos de sistemas: basados en reglas previamente establecidas, basados en casos y basados en redes bayesianas. En cada uno de ellos, la solución a un problema dado se obtiene, respectivamente: aplicando reglas heurísticas apoyadas generalmente en lógica difusa para su evaluación y aplicación; aplicando el razonamiento basado en casos, donde la solución a un problema similar planteado con anterioridad se adapta al nuevo problema; y aplicando redes bayesianas, basadas en estadística y el Teorema de Bayes.

Entre sus ventajas se puede destacar la mejora de la productividad al resolver problemas rápidamente, ahorrando tiempo y dinero al que se necesitaría si se tratase de un experto humano; los valiosos conocimientos del especialista se guardan y se difunden, de forma que no se pierden aunque éste desaparezca; se obtienen soluciones más fiables gracias al tratamiento automático de los datos; tienen gran flexibilidad debido a la separación entre la base de conocimiento y el mecanismo de inferencia, lo que se traduce en una mejor modularidad y legibilidad del conocimiento.

Pese a estas ventajas, también existen algunos inconvenientes, ya que el conocimiento humano es complejo de extraer y, a veces, difícil de representar. Además, las estrategias de razonamiento de los motores de inferencia suelen estar programadas procedimentalmente y se adaptan mal a ciertas circunstancias. Tienen cierta limitación para tratar problemas con información incompleta, es por ello que, durante el desarrollo del sistema, se necesita una realimentación que haga posible rectificar, si es necesario, las reglas de acuerdo a la experiencia.

2.3.1. Sistemas expertos basados en reglas

Los sistemas expertos se aplican a diferentes campos, entre ellos, muy frecuentemente a la medicina, ámbito en el que se centra este trabajo. En éstos, es habitual utilizar los sistemas basados en reglas ya que tratan situaciones complejas gobernadas por decisiones deterministas. Su base de conocimiento contiene las variables y el conjunto de reglas que definen el problema, definidas por el experto, y el motor de inferencia obtiene las conclusiones aplicando lógica a estas reglas.

Pueden utilizarse dos mecanismos básicos de inferencia: encadenamiento hacia adelante (*forward chaining*) que comienza con una evidencia inicial en una determinada situación y se dirige hacia la obtención de una solución, o encadenamiento hacia atrás (*backward chaining*) que parte de una hipótesis sobre las posibles soluciones y vuelve hacia atrás para encontrar una evidencia existente (o una deducción de una evidencia existente) que apoye una hipótesis en particular.

En este trabajo se va a utilizar CLIPS (*C Language Integrated Production System*) como herramienta para construir el sistema, ya que proporciona encadenamiento hacia adelante (*forward*), lo cual permite simular el razonamiento deductivo habitual de un experto [16]. Algunas de las características que han favorecido a la elección de CLIPS son:

- Permite manejar una amplia variedad de conocimiento, soportando tres estilos de programación: procedural, orientado a objetos y basado en reglas.
- Es portable y puede ser ejecutado en cualquier sistema con un compilador ANSI de C, o un compilador de C++.
- Puede ser embebido en código imperativo, invocado como una sub-rutina, e integrado con lenguajes como C, Java, FORTRAN y otros.
- Es un software de dominio público.

El sistema experto MYCIN

Actualmente, el sistema MYCIN [17] es considerado uno de los sistemas expertos de más prestigio. Fue desarrollado en la Universidad de Stanford como sistema de consulta para el diagnóstico y tratamiento de enfermedades infecciosas. Es, sin duda, el sistema experto más famoso, el que ha dado lugar a más proyectos derivados y el que ha marcado el paradigma de todos los sistemas expertos de la actualidad.

MYCIN representa el razonamiento experto como una serie de reglas condición - conclusión, que relacionan los datos del paciente con hipótesis de infección a la vez que estiman la certeza de cada regla. Funciona con encadenamiento hacia atrás a partir de un diagnóstico como hipótesis y, con reglas que estiman los *factores de certeza* de las conclusiones propagados desde la evidencia, comprueba si ésta apoya el diagnóstico. Si no hay información suficiente para comprobar la hipótesis, pedirá al experto datos adicionales y evaluará de forma exhaustiva todas las hipótesis. Cuando ha finalizado, MYCIN da los tratamientos para todas aquellos diagnósticos que han alcanzado un alto valor de certeza.

El *factor de certeza* o *factor de certidumbre* (CF) de cada regla se define como el grado de creencia en la hipótesis dada la evidencia. Más concretamente, como la diferencia entre la creencia a favor (*measure of belief*, MB) y la creencia en contra (*measure of disbelief*, MD):

$$MB(H, E) = \begin{cases} \frac{\mathbb{P}(H|E) - \mathbb{P}(H)}{1 - \mathbb{P}(H)} & \text{si } \mathbb{P}(H|E) \geq \mathbb{P}(H) \\ 0 & \text{si } \mathbb{P}(H|E) < \mathbb{P}(H) \end{cases} \quad (2.1)$$

$$MD(H, E) = \begin{cases} 0 & \text{si } \mathbb{P}(H|E) \geq \mathbb{P}(H) \\ \frac{\mathbb{P}(H) - \mathbb{P}(H|E)}{\mathbb{P}(H)} & \text{si } \mathbb{P}(H|E) < \mathbb{P}(H) \end{cases} \quad (2.2)$$

$$CF(H, E) = MB(H, E) - MD(H, E) = \begin{cases} \frac{\mathbb{P}(H|E) - \mathbb{P}(H)}{1 - \mathbb{P}(H)} & \text{si } \mathbb{P}(H|E) \geq \mathbb{P}(H) \\ \frac{\mathbb{P}(H|E) - \mathbb{P}(H)}{\mathbb{P}(H)} & \text{si } \mathbb{P}(H|E) < \mathbb{P}(H) \end{cases} \quad (2.3)$$

Siendo H la hipótesis y E la evidencia que conforman la regla. MB es una medida proporcional del aumento de la credibilidad (no de la credibilidad absoluta); exactamente, es una medida

de la disminución proporcional de la falta de credibilidad, $1 - \mathbb{P}(H)$. Del mismo modo, MD es una medida proporcional del aumento de la creencia en contra de H ; es decir, es una medida de la disminución proporcional de $\mathbb{P}(H)$. Además, los factores de certeza cumplen: $0 \leq MB \leq 1$, $0 \leq MD \leq 1$ y, por tanto, $-1 \leq CF \leq 1$.

Cuando se tiene una regla simple, la confirmación de la evidencia E conlleva la confirmación de H . Sin embargo, cuando hay una red de reglas, esa evidencia se propaga mediante la aplicación repetida de combinaciones. Es el caso de los sistemas expertos, donde las reglas suelen ser una combinación de reglas cuyo antecedente está formado por una cláusula compleja. La forma en que MYCIN [18] trata este tipo de reglas consiste en tomar el opuesto del factor de certeza para la negación, el mínimo de los factores de certeza para la conjunción y el máximo para la disyunción, como sigue:

$$CF(\neg E) = -CF(E) \quad (2.4)$$

$$CF(E_1 \wedge E_2) = \min [CF(E_1), CF(E_2)] \quad (2.5)$$

$$CF(E_1 \vee E_2) = \max [CF(E_1), CF(E_2)] \quad (2.6)$$

2.3.2. Ontologías

Las ontologías informáticas son, en esencia, clasificaciones conceptuales utilizadas como medio para categorizar y agrupar información en conceptos y reglas de correspondencia. Se trata de una especificación explícita y formal de una conceptualización compartida [19] que cumple con las siguientes características:

- Explícita: compuesta de conceptos, propiedades, relaciones, funciones, axiomas y restricciones.
- Formal: puede ser interpretada por máquinas.
- Conceptualizada: modelo abstracto del dominio a representar.
- Compartida: acordada por grupos de expertos.

Esta estructura se ha utilizado durante años en inteligencia artificial y sistemas expertos, posteriormente en agentes inteligentes y web semántica y actualmente en cualquier aplicación que requiera establecer un modelo de información, como es este caso. A efectos prácticos, estas estructuras semánticas tienen las ventajas de trabajar con los conceptos y sus relaciones con otros conceptos lo que da lugar a una mejor organización de la información y, por tanto, una optimización de su definición.

Para construir una ontología es necesario seleccionar el lenguaje de especificación. OWL (*Ontology Web Language*) es el más aceptado como estándar para representar y compartir

conocimiento en la web semántica [20]. Este lenguaje está basado en RDF (*Resource Description Framework*) y codificado en el metalenguaje XML.

Además, en la actualidad, se pueden aprovechar desarrollos realizados en el ámbito de la web semántica como la edición de ontologías donde se utiliza Protégé [21]. Precisamente es esta herramienta la que se utilizará en este proyecto para verificar la coherencia del sistema experto, la cual está implementada en Java y dispone de una interfaz gráfica que permite definir todos los conceptos como clases y sus relaciones como conexiones entre clases, permitiendo almacenar toda la definición en lenguaje OWL.

2.4. Explicación en clasificación

El aprendizaje automático está adquiriendo cada vez más importancia en diversas ramas de la ciencia, la tecnología o los negocios. En ciencias de la computación el aprendizaje automático es una rama de la inteligencia artificial cuyo objetivo es desarrollar técnicas que permitan a las computadoras aprender. De forma más concreta, el objetivo principal del aprendizaje automático es la creación de un modelo capaz de dar una respuesta satisfactoria cuando se introduce información al mismo. En función de la realimentación que reciba el modelo se distinguen varios paradigmas. Este proyecto se enmarca en el ámbito del aprendizaje supervisado, que se caracteriza por el conocimiento de cuáles son las salidas esperadas para cierto conjunto de datos de entrada etiquetados.

Sin embargo, la salida que ofrece el modelo puede no ser satisfactoria o quizás resulte difícil de interpretar. En ciertos campos de aplicación, como la medicina, el profesional que emplea modelos de aprendizaje automático demanda sistemas con capacidad de explicar sus predicciones. Esto es, si cuando el sistema emite una predicción, la acompaña de información adicional que la justifique, el experto podrá valorar adecuadamente la decisión tomada, en lugar de solamente aceptar o rechazar, a ciegas, el resultado. De este modo, puede valorar más información que le apoye en la decisión final y, quizás, adquirir nuevos conocimientos sobre el problema abordado.

Existen diversas técnicas que tratan de dotar de mayor transparencia a los modelos, especialmente aquellos considerados tradicionalmente como de *caja negra*, por ejemplo las *redes neuronales*. En [5] se estudian diferentes aproximaciones propuestas en la literatura, pero con algunas limitaciones que, finalmente, se solucionan con una propuesta de Strumbelj y Kononenko en [22] basada en la teoría de juegos cooperativos.

Un *juego cooperativo* es un juego en el cual dos o más jugadores no compiten, sino que se esfuerzan por conseguir el mismo objetivo y, por lo tanto, ganan o pierden en conjunto. La idea básica consiste en identificar los jugadores con atributos (o los valores que éstos toman) y las coaliciones con subconjuntos de atributos. De este modo se pueden emplear los resultados conocidos de la teoría de juegos cooperativos para determinar la influencia que tienen los atributos

sobre las predicciones de un modelo. Formalmente, se define como sigue.

Definición 1 (Juego Cooperativo)

Un juego cooperativo es una tupla $\langle \mathbb{A}, \nu \rangle$ donde $\mathbb{A} = \{1, \dots, a\}$ es un conjunto de a jugadores (atributos) y $\nu : \mathcal{P}(\mathbb{A}) \rightarrow \mathbb{R}$ (con $\nu(\emptyset) = 0$) es el valor de la recompensa asignada a cada posible coalición de jugadores (subconjunto de atributos). Donde $\mathcal{P}(\mathbb{A})$ denota el conjunto de todos los subconjuntos de \mathbb{A} .

En base a esta definición, el objetivo es repartir la recompensa total ($\nu(\mathbb{A})$, gran coalición) de forma «justa» entre todos los jugadores, teniendo en cuenta la influencia de cada jugador en todas las posibles sub-coaliciones en las que participe. El resultado de ello es un vector $\phi(\nu) = (\phi_1, \dots, \phi_a) \in \mathbb{R}^a$ donde ϕ_j es la recompensa asignada al jugador j . Dado un juego cooperativo, existen muchas formas de asignar $\phi(\nu)$ pero solamente se consideran «justas» las que cumplan las condiciones de *eficiencia*, *simetría*, *irrelevancia* y *aditividad*, descritas en [5].

En [23], Shapley demuestra que solamente existe una manera de asignar $\phi(\nu)$ que cumpla las condiciones anteriores. Dicha solución es lo que se denomina *Valor de Shapley*, que es un concepto básico de la teoría de juegos cooperativos. El siguiente teorema recoge dicho resultado.

Teorema 1 (Valor de Shapley)

Dado el juego cooperativo $\langle \mathbb{A}, \nu \rangle$, existe una única solución que satisface las cuatro condiciones citadas anteriormente:

$$\phi(\nu) = Sh_j(\nu) = \sum_{S \subseteq \mathbb{A} \setminus \{j\}} \frac{(a - s - 1)!s!}{a!} (\nu(S \cup \{j\}) - \nu(S)), \quad (2.7)$$

con $a = |\mathbb{A}|$, $s = |S|$, $j = 1, \dots, a$.

En la solución propuesta por Strumbelj y Kononenko en [22] se asigna un peso numérico al valor que toma cada atributo de una instancia, en función de la contribución de dicho valor a la predicción de una clase determinada. Así, teniendo en cuenta las contribuciones de los valores de cualquier subconjunto posible de atributos, también se puede valorar la influencia que tienen dichos atributos en las predicciones. Para ello, lo único que se requiere es un modelo capaz de proporcionar probabilidades a posteriori de las clases. La idea descrita en la Definición 2 consiste en descomponer la predicción para cualquier subconjunto posible de atributos S y analizar la variación de la predicción de la instancia I cuando se conoce el valor de todo el subconjunto y cuando se suponen desconocidos los valores de S ($\mathbb{A} \setminus S$).

Definición 2 (Diferencia en la predicción)

Sea $S \subseteq \mathbb{A}$ un subconjunto de atributos. Suponiendo conocidos los valores de los atributos de S y desconocidos los de $\mathbb{A} \setminus S$, se define la diferencia en la predicción de la clase C_k para una instancia I como

$$\Delta(S) = \frac{1}{|\mathcal{A}_{\mathbb{A} \setminus S}|} \sum_{J \in \mathcal{A}_{\mathbb{A} \setminus S}} \hat{f}_k(\lambda(I, J, S)) - \frac{1}{|\mathcal{A}_{\mathbb{A}}|} \sum_{J \in \mathcal{A}_{\mathbb{A}}} \hat{f}_k(J), \quad (2.8)$$

donde :

- $\mathcal{A}_{\mathbb{A}}$ representa el espacio completo de atributos.
- $\mathcal{A}_{\mathbb{A} \setminus S}$ es el subespacio de $\mathcal{A}_{\mathbb{A}}$ correspondiente a fijar los valores de los atributos de S y explorar todas las combinaciones de valores de los atributos de $\mathbb{A} \setminus S$.
- λ construye una nueva instancia, combinando la original (I) con otra tomada de $\mathcal{A}_{\mathbb{A} \setminus S}(J)$. Los atributos de S toman el valor que tienen en I , mientras que los que no están en S toman el valor que tienen en J .

$$\lambda(I, J, S) = (\alpha_1, \alpha_2, \dots, \alpha_a), \quad \alpha_j = \begin{cases} \psi(A_j, I) & j \in S \\ \psi(A_j, J) & j \notin S \end{cases}$$

- $\hat{f}_k(I) = \mathbb{P}(C_k|I)$ representa la predicción correspondiente a la clase C_k , ofrecida por el modelo para la instancia I .

La definición original del *Valor de Shapley* (Definición 1) explora todos los subconjuntos de atributos, haciendo que su cálculo tenga alto coste computacional. Por ello, se incluye una formulación alternativa (Proposición 1), basada en permutaciones, las cuales son sencillas de manejar por los algoritmos de computación.

Proposición 1 (Formulación alternativa del Valor de Shapley)

$$Sh_j(\Delta) = \frac{1}{a!} \sum_{\pi \in \Pi(\mathbb{A})} (\Delta(Pre^j(\pi) \cup \{j\}) - \Delta(Pre^j(\pi))), \quad j = 1, \dots, a. \quad (2.9)$$

Siendo

- $\Pi(\mathbb{A})$ el conjunto de las permutaciones del conjunto $\mathbb{A} = \{1, \dots, a\}$.
- $Pre^j(\pi)$ el conjunto de elementos (atributos) que preceden a j en la permutación π .

Teniendo en cuenta que $|\mathcal{A}_{\mathbb{A}}| = |\mathcal{A}_{\mathbb{A} \setminus S}| \cdot |\mathcal{A}_S|$, se puede reescribir $\Delta(S)$ en la Definición 2.

$$\Delta(S) = \frac{1}{|\mathcal{A}_{\mathbb{A}}|} \sum_{J \in \mathcal{A}_{\mathbb{A}}} (\hat{f}_k(\lambda(I, J, S)) - \hat{f}_k(J)),$$

donde $\hat{f}_k(\lambda(I, J, S))$ aparece $|\mathcal{A}_S|$ veces.

Introduciendo este resultado en la Proposición 1 se obtiene la contribución teórica que el método de explicación otorga a cada atributo.

$$\varphi_j(\Delta) = \frac{1}{a!|\mathcal{A}_{\mathbb{A}}|} \sum_{\pi \in \Pi(\mathbb{A})} \sum_{J \in \mathcal{A}_{\mathbb{A}}} \left[\hat{f}_k(\lambda(I, J, Pre^j(\pi) \cup \{j\})) - \hat{f}_k(\lambda(I, J, Pre^j(\pi))) \right]. \quad (2.10)$$

Además, si se define la variable aleatoria $X : \Pi(\mathbb{A}) \times \mathcal{A}_{\mathbb{A}} \rightarrow \mathbb{R}$ como:

$$X_{\pi \in \Pi(\mathbb{A}), J \in \mathcal{A}_{\mathbb{A}}} = \hat{f}_k(\lambda(I, J, Pre^j(\pi) \cup \{j\})) - \hat{f}_k(\lambda(I, J, Pre^j(\pi))).$$

Se observa que $\varphi_j = \mathbb{E}[X_{\pi, J}]$ y se define $\hat{\varphi}_j = \frac{1}{m} \sum_{l=1}^m X_l$, donde $\{X_l\}_{l=1}^m$ es un conjunto de variables aleatorias independientes e idénticamente distribuidas (v.a.i.i.d) con distribución igual

a la variable aleatoria X y m es el cardinal de dicho conjunto. Entonces, por el Teorema del Límite Central, $\hat{\varphi}_j$ tiende a una distribución normal con media φ_j y varianza $\frac{\sigma_j^2}{m}$. Por tanto, $\hat{\varphi}_j$ es un estimador de φ_j no sesgado ($\mathbb{E}(\hat{\varphi}_j) = \varphi_j$) y consistente ($\frac{\sigma_j^2}{m} \rightarrow 0$).

Basándose en todas las deducciones anteriores, en [5] se propone el Algoritmo 1 para estimar las contribuciones de cada atributo.

Algoritmo 1 Estimación de la contribución de un atributo mediante teoría de juegos cooperativos

Entrada: Modelo entrenado para la clase $C_k(\hat{f}_k)$, instancia I .

Salida: Valor estimado de la contribución del atributo A_j para clasificar I como perteneciente a la clase C_k .

```

1: Fijar  $m$ .
2:  $\varphi_j \leftarrow 0$ .
3: for  $l = 1 \rightarrow m$  do
4:   Seleccionar aleatoriamente una permutación  $\pi \in \Pi(\mathbb{A})$ .
5:   Seleccionar aleatoriamente una instancia  $J \in \mathcal{A}_{\mathbb{A}}$ .
6:    $\nu_1 \leftarrow \hat{f}_k(\lambda(I, J, Pre^j(\pi) \cup \{j\}))$ .
7:    $\nu_2 \leftarrow \hat{f}_k(\lambda(I, J, Pre^j(\pi)))$ .
8:    $\varphi_j \leftarrow \varphi_j + (\nu_1 - \nu_2)$ .
9: end for
10:  $\varphi_j \leftarrow \frac{\varphi_j}{m}$ .
11: return  $\varphi_j$ 

```

El valor de m se fija en función del error máximo que se desee tolerar (ϵ). Al tratarse de un muestreo aleatorio, no se puede garantizar que el 100 % de las contribuciones $\hat{\varphi}_j$ calculadas mediante el algoritmo anterior tengan un error menor que ϵ . Pero, gracias al Teorema del Límite Central, se puede escoger una probabilidad $1 - \alpha$ con la que se cumpla, como se detalla a continuación:

$$\begin{aligned} \mathbb{P}(|\hat{\varphi}_j - \varphi_j| < \epsilon) &= 1 - \alpha \Rightarrow \mathbb{P}\left(\left|\frac{1}{m} \left(\sum_{l=1}^m X_l\right) - \varphi_j\right| < \epsilon\right) = 1 - \alpha \\ &\Rightarrow \mathbb{P}\left(\frac{|\sum_{l=1}^m X_l - m \cdot \varphi_j|}{\sigma_j \sqrt{m}} < \frac{m \cdot \epsilon}{\sigma_j \sqrt{m}}\right) = 1 - \alpha. \end{aligned}$$

Tomando z tal que $\mathbb{P}(\mathcal{N}(0, 1) \leq z) = 1 - \frac{\alpha}{2}$ y sustituyendo en la expresión anterior se tiene:

$$z = \frac{m \cdot \epsilon}{\sigma_j \sqrt{m}} \Rightarrow m = \frac{z^2 \cdot \sigma_j^2}{\epsilon^2}.$$

A priori no se conoce el valor de σ_j pero se toma la cota superior: $\sigma_j = 1$, suponiendo que X estuviese uniformemente distribuida entre sus dos extremos (-1 y 1). De esta forma, por ejemplo, para que el 99 % de las contribuciones tengan un error menor que 0,01, z es 2,57 y $m \approx 65.000$ iteraciones.

En la Sección 6.5 se presentan algunos de los resultados obtenidos al implementar este método, mediante el cual aumenta la transparencia del modelo y facilita la interpretación de sus predicciones.

3

Sistema de recuperación de información

Uno de los principales hitos del desarrollo del proyecto es obtener la información recogida en campos de texto libre de la Historia Clínica Electrónica y transformarla en datos estructurados y normalizados. Para ello, será necesario diseñar una herramienta que permita llevar a cabo un proceso de recuperación de información (IR) dentro del ámbito médico.

Concretamente, se van a extraer los datos haciendo un preprocesado con el objetivo de eliminar la información redundante e irrelevante que pudiese existir. Después de este proceso de extracción, es necesario realizar una validación de la información recopilada. Para este fin, se diseña un sistema de validación que permita a los expertos acceder tanto a la Historia Clínica como a los datos extraídos ya estructurados, y así poder validar la implementación de este sistema, entre otras opciones.

En esta sección se explica, en primer lugar, el conjunto de datos inicial y se analizan los atributos que lo compondrán, incluyendo la fase de preprocesamiento. Posteriormente, se presenta el sistema desarrollado para extraer la información clínica y, por último, el sistema de validación implementado.

3.1. Conjunto de datos

Como se ha detallado en la introducción de este proyecto, el principal objetivo es desarrollar un modelo predictivo del éxito de un procedimiento de inducción del trabajo de parto. El punto de partida es, por tanto, la recopilación de todos los datos que componen el conjunto, extraídos de los registros de la Historia Clínica Electrónica en la plataforma Selene [15] del Hospital Universitario de Fuenlabrada [24].

En concreto, se recogen todos los datos de episodios asistenciales de embarazadas en un periodo de más de 5 años, cuyo informe de parto ha sido cerrado después del 1 de Enero de 2010 hasta el 31 de Julio de 2015. Estos datos incluyen todos los objetos clínicos registrados en la Historia Clínica Electrónica ya enumerados en el estado del arte: *informes*, *notas*, *formularios* y *peticiones*. Se tratan de datos, la mayoría en formato de texto libre, categorizados por ciertas variables (atributos) definidas por el experto, que deberán ser extraídas de manera automática por el sistema de recuperación de información.

Este trabajo está enmarcado en el proyecto INDUCCESS (INDUCtion y sucCESS), cuya propuesta ha sido revisada y aprobada por el Comité Ético de Investigación Clínica y por la Agencia Española del Medicamento, asegurando así que se ajusta a la legislación vigente y contiene los principios éticos aplicables a la investigación en seres humanos. En la realización de éste participan como investigadores principales del hospital el responsable de Sistemas de Información, Luis Lechuga, y la responsable del Servicio de Ginecología y Obstetricia, doctora María Teulón. El Instituto de Ingeniería del Conocimiento (IIC) [25], entidad a la que pertenezco, se incorpora como parte del proyecto y tiene acceso a los datos, totalmente anonimizados, en cumplimiento de la L.O.P.D. (Ley Orgánica de Protección de Datos de Carácter Personal).

3.1.1. Preproceso de datos

En el proceso de extracción de datos se incluyen todos los embarazos atendidos en el Hospital de Fuenlabrada que tienen asignado un proceso asistencial en Selene, siendo excluidos aquellos con parto en otro centro y los que tienen un proceso de alta de parto con fetos con edad gestacional inferior a 23 semanas (límite actual establecido de la viabilidad fetal). Todos esos datos incluidos en el estudio pueden estar incompletos, contener información innecesaria o ambigua, sufrir perturbaciones a causa del ruido o presentar cualquier otra dificultad que afecte al rendimiento de los modelos. Por ello, es necesario preprocesarlos y así solventar futuros inconvenientes.

En primer lugar, se eliminan del conjunto aquellos objetos clínicos en los que no aparezca especificado el tipo (*informes*, *notas*, *formularios* o *peticiones*) al que pertenecen. Este tratamiento es imprescindible en el proceso de extracción, ya que cada atributo se trata de manera particular y se extrae su información de un subconjunto de objetos clínicos definido previamente.

En segundo lugar, existen ciertos atributos como son el *índice de Bishop al ingreso*, la *edad gestacional*, el *peso fetal estimado* o el *peso de los neonatos* que requieren un tratamiento especial. Concretamente, la información del *índice de Bishop al ingreso* no aparece especificada en los datos, sino que es necesario inferir su valor a partir de otras cuatro variables (*posición*, *consistencia*, *borramiento*, *dilatación cérvix*) que lo componen. En el caso de la *edad gestacional* o los pesos, en cambio, el tratamiento no es tan complejo, y solamente es necesario realizar una conversión en sus unidades de medida.

3.1.2. Diseño del modelo de datos

Como se ha mencionado en el subapartado anterior, el espacio de atributos que compone el modelo predictivo es fijado previamente por el personal experto en este ámbito. En la Figura 3.1 se presentan los nombres de los 46 atributos claramente separados en dos categorías, relegando al Anexo A la descripción de cada uno. En azul aparecen aquellas variables que, en el estado del arte, se denominan predictoras (participan en la decisión de inducir el trabajo de parto) y, en negro, aquellas de resultado, es decir, no aportan nada en la decisión y de ellas se puede deducir directamente la clase.

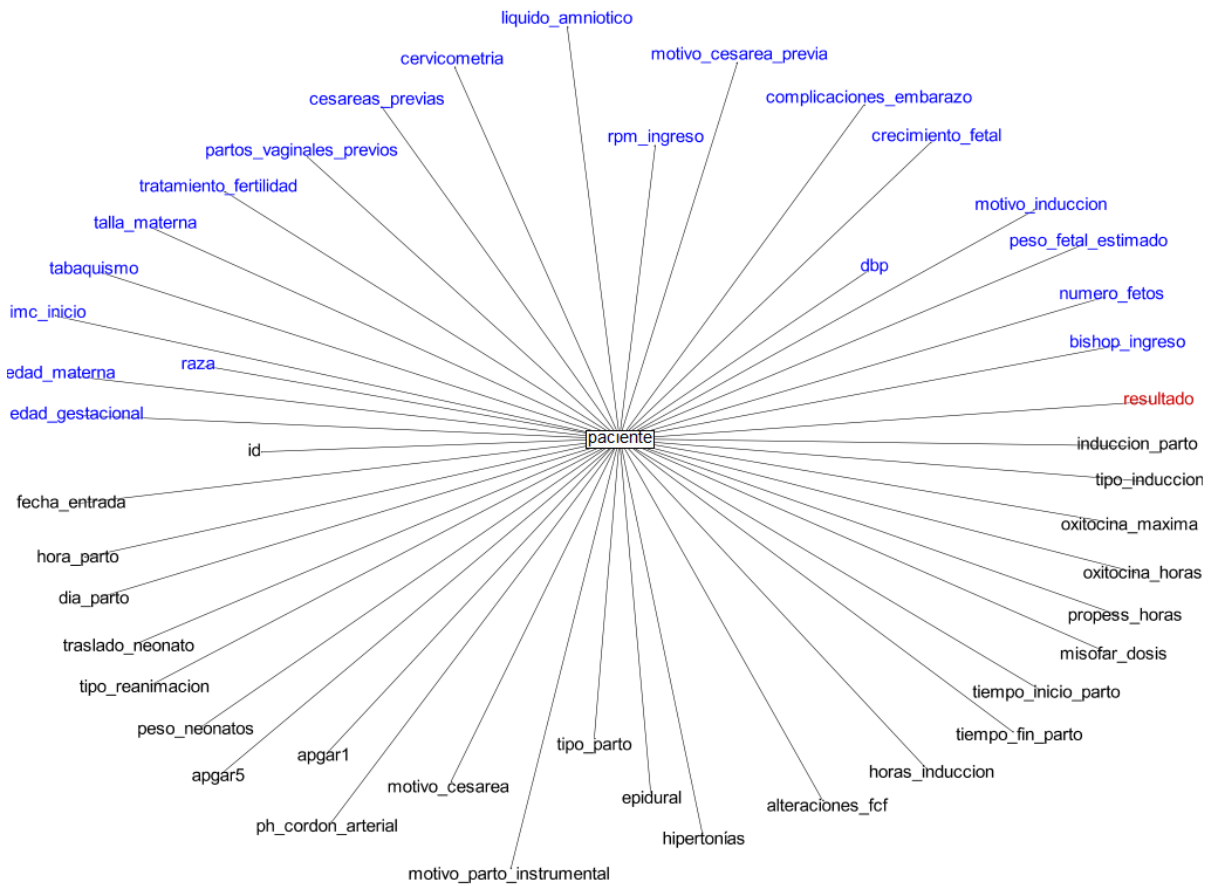


Figura 3.1: Grafo de atributos de un paciente. En azul aparecen los atributos que participan en la decisión de inducir el trabajo de parto, en negro aquellos que no aportan nada en la decisión y en rojo la variable objetivo (clase).

Entre estos atributos numéricos y nominales, se encuentra marcada en rojo la variable objetivo, *resultado*, la cual toma tres posibles valores: *No inducción*, *Inducción* o *Cesárea*. Serán, precisamente, estos valores las posibles salidas (clases) que devuelvan los modelos predictivos. Como se verá más adelante, la proporción de estas clases no está balanceada, es decir, existen más instancias asignadas a una clase que al resto, lo que puede suponer un problema para cierto tipo de clasificadores.

3.2. Diseño del sistema IR

La recuperación de información del conjunto de datos inicial unida al proceso de estructuración y normalización son de las tareas esenciales y, a su vez, unas de las más complejas de este proyecto. Sin dicha información ya estructurada no sería posible generar los modelos predictivos ni construir la herramienta de validación.

Para dar soporte a ello, ha sido necesario diseñar un robusto sistema de extracción de información. Este sistema está compuesto por una serie de módulos y pasos a seguir, los cuales se muestran en la Figura 3.2.

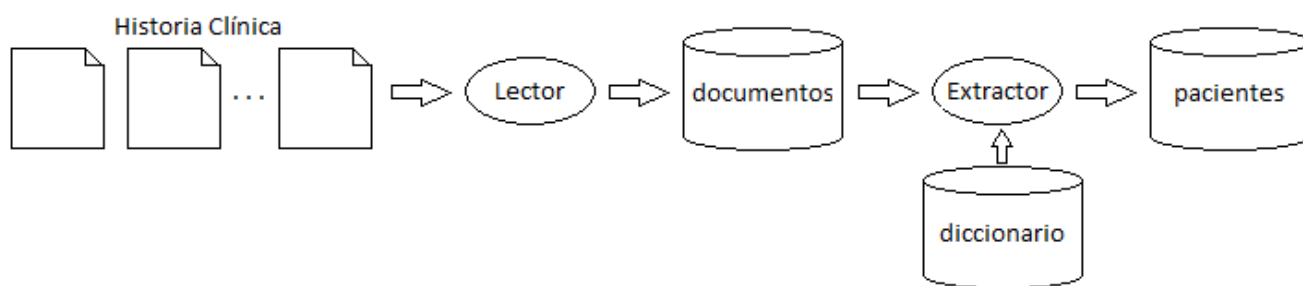


Figura 3.2: Proceso seguido en el sistema IR.

A continuación se describen cada una de las partes de este diagrama.

Historia Clínica

El conjunto de datos proporcionado por el hospital ha sido exportado en una serie de ficheros de texto plano de Selene. Cada uno de ellos, además de la información anonimizada de los pacientes, incluye metadatos y el tipo de objeto clínico, esto es, *informe*, *nota*, *formulario* o *petición*, lo que permite agruparlos por categorías.

Lector y documentos

Una vez identificados los tipos de objeto de la Historia Clínica de cada paciente, es necesario el diseño e implementación de una base de datos estructurada para que el sistema de recuperación pueda acceder a dicha información de una forma rápida y eficiente. Para ello, se ha utilizado el sistema gestor de bases de datos relacionales (RDBMS) PostgreSQL¹ generando tres tablas: *documentos*, *diccionario* y *pacientes*.

Los metadatos de los ficheros exportados de Selene son tratados antes de ser introducidos en la base de datos, mientras que el contenido del propio fichero se almacena íntegramente. Para este fin se ha diseñado un programa lector programado en Awk [26]. Algunos de los

¹Se puede consultar más información en <http://www.postgresql.org.es/>

tratamientos efectuados son la homogeneidad de capitalización o la eliminación de caracteres especiales, además de realizar un pequeño filtro de los metadatos realmente relevantes en el proyecto. Todos estos datos (metadatos y contenido) son almacenados junto al identificador del paciente en la tabla *documentos*.

Al finalizar el proceso, tras eliminar aquellos datos descritos en el subapartado 3.1.1, esta tabla consta de 1.403.040 objetos clínicos, de los cuales 3.509 son informes, 399.646 notas no estructuradas, 764.783 formularios y 235.102 peticiones de pruebas.

Extractor y diccionario

Una vez almacenada toda la información en la base de datos, se procede a realizar la extracción de atributos de los pacientes.

Para ello, se ha implementado, por un lado, una tabla *diccionario*, alimentada por un panel de expertos en este ámbito según su experiencia y conocimiento y, por otro lado, un programa extractor utilizando el lenguaje de programación Ruby [27].

El *diccionario* se actualiza tras cada validación del experto, ya que la extracción debe ser un proceso iterativo y no siempre se logran buenos resultados a la primera. El objetivo de esta tabla es guiar al programa que realiza la extracción de atributos y, para ello, contiene información para cada variable de paciente y para cada nombre de objeto clínico, el tipo de dato a extraer (*DATO* o *TEXTO*) y su prioridad. Si se trata de un atributo *DATO*, el extractor guardaría el valor directamente; pero si, por el contrario, se trata de *TEXTO*, se indica además la palabra o cadena a buscar y el valor a insertar en la variable del paciente. De este modo los atributos de la tabla son: *nom_column*, *nom_document*, *nom_variable*, *prioridad*, *tipo*, *clave* y *resultado*.

En cuanto al programa implementado, éste se alimenta de los datos almacenados en la tabla *documentos* y de dicho *diccionario*. Como ya se ha explicado previamente, el contenido de la Historia Clínica se almacena en la base de datos sin modificar, por ello ahora es necesario realizar un tratamiento previo de la información. Este tratamiento consiste en una limpieza del vocabulario con las siguientes tareas de procesamiento:

- Extracción de vocabulario

El primer paso a realizar es la extracción de todos los términos mediante la separación del texto de cada contenido en palabras, utilizando como separadores los espacios o los signos de puntuación. También se realiza una conversión de los caracteres a mayúsculas y se sustituyen las letras acentuadas por su correspondiente letra sin acento. Este paso se realiza a la par que el proceso de *stemming*, descrito a continuación.

- *Stemming*

El objetivo del *stemming* es reducir inflexiones y derivaciones a un lema común. La idea es quedarse con significados e ignorar diferencias irrelevantes. De esta forma el programa

solamente debe buscar el lexema de las palabras y se evitan problemas con el género y número de las palabras.

■ Eliminación de *stopwords*

Se eliminan las partículas funcionales con poco significado propio como artículos, preposiciones o conjunciones, además de palabras muy frecuentes. El objetivo de este tratamiento es eliminar ruido y así reducir el tamaño del contenido a parsear.

Una vez llevado a cabo el preprocesado del texto se procede a buscar cada uno de los atributos de los pacientes. Para ello el programa implementado basa su funcionamiento en el Algoritmo 2.

Algoritmo 2 Proceso de extracción de atributos

Entrada: Tabla *documentos* y *diccionario* de la *BD*.

Salida: Tabla *pacientes* de la *BD*.

```

1: for  $paciente_i$  in documentos do
2:   for  $A_j$  in diccionario do
3:     Extraer todos los nom_document ordenados por prioridad para  $paciente_i$  y  $A_j$ .
4:     for doc in nom_document do
5:       if doc.tipo == DATO then
6:         Buscar nom_variable en doc.
7:          $A_{ji} \leftarrow valor$ .
8:       else, if doc.tipo == TEXTO then
9:         Buscar nom_variable y clave en doc.
10:         $A_{ji} \leftarrow resultado$ .
11:      end if
12:    end for
13:    if  $A_{ji} \neq null$  then
14:      Añadir  $A_{ji}$  al  $paciente_i$ .
15:    else
16:      Dejar vacío  $A_{ji}$  en  $paciente_i$ .
17:    end if
18:  end for
19:  Tratamiento especial de variables.
20:   $BD \leftarrow paciente_i$ 
21: end for
22: return BD

```

Pacientes

Para finalizar el proceso de recuperación de información expuesto, se insertan los datos de cada paciente ya estructurados por atributos en la base de datos. Para ello se ha creado la tabla *pacientes*, en la que se estructuran todos los atributos relativos a cada paciente. Estos atributos, tanto predictores como de resultado, se incluyen en la Tabla 3.1. Como resultado del proceso de extracción llevado a cabo, esta tabla contiene 10.487 pacientes con los que trabajar.

Tabla 3.1: Atributos de la tabla *pacientes* de la base de datos.

<i>id</i>	<i>dbp</i>	<i>hipertonias</i>
<i>edad_gestacional</i>	<i>motivo_induccion</i>	<i>epidural</i>
<i>edad_materna</i>	<i>peso_fetal_estimado</i>	<i>tipo_parto</i>
<i>raza</i>	<i>bishop_ingreso</i>	<i>motivo_parto_instrumental</i>
<i>imc_inicio</i>	<i>numero_fetos</i>	<i>motivo_cesarea</i>
<i>tabaquismo</i>	<i>inducción_parto</i>	<i>ph_cordon_arterial</i>
<i>talla_materna</i>	<i>tipo_induccion</i>	<i>apgar1</i>
<i>tratamiento_fertilidad</i>	<i>oxitocina_maxima</i>	<i>apgar5</i>
<i>partos_vaginales_previos</i>	<i>oxitocina_horas</i>	<i>peso_neonatos</i>
<i>cesareas_previas</i>	<i>propess_horas</i>	<i>tipo_reanimacion</i>
<i>cervicometria</i>	<i>misofar_dosis</i>	<i>traslado_neonato</i>
<i>liquido_amniotico</i>	<i>tiempo_inicio_parto</i>	<i>dia_parto</i>
<i>rpm_ingreso</i>	<i>tiempo_fin_parto</i>	<i>hora_parto</i>
<i>motivo_cesarea_previa</i>	<i>horas_induccion</i>	<i>fecha_entrada</i>
<i>complicaciones_embarazo</i>	<i>alteraciones_fcf</i>	<i>resultado</i>
<i>crecimiento_fetal</i>		

Por último, se debe destacar que todo el proceso descrito previamente ha sido un trabajo muy costoso computacionalmente.

En una primera versión, el tiempo de ejecución fue de aproximadamente 15 días, tiempo inaceptable, teniendo en cuenta que este proceso debe ser repetido tras cada validación del experto. Como solución a esto, se implementó una nueva versión del Algoritmo 2 paralelizándolo mediante hilos. Su funcionamiento es el mismo con la única diferencia de que se lanzan de manera paralela 35 hilos y cada uno se encarga de la extracción de un subconjunto de pacientes. En esta segunda versión, se redujo 3 veces el tiempo y, ahora se necesitan 5 días para completar el proceso. Puesto que sigue siendo un tiempo demasiado elevado, se buscaron otras soluciones posibles y, la que finalmente se añadió, es la optimización del índice creado en la tabla *documentos*, ya que se pudo observar que el coste computacional recaía en la lectura de la base de datos. Con este nuevo índice mejora la velocidad de las operaciones, por medio de un identificador único de cada fila de la tabla, permitiendo un rápido acceso a los registros. Así, el tiempo se vio reducido considerablemente a 3 horas aproximadamente.

3.3. Sistema de validación

El proceso de extracción de atributos es un proceso largo y tedioso, que necesita la colaboración del experto para validar que la extracción es correcta. Tras cada validación, el *diccionario* se actualiza con el objetivo de lograr los mejores resultados. Por ello, es necesario disponer de un sistema de validación donde varios expertos puedan acceder de manera paralela tanto a la Historia Clínica como a los datos estructurados de los pacientes. Se ha diseñado, por tanto, una

aplicación web con el objetivo de facilitar a los expertos dicha validación y la visualización de la información clínica. El acceso a la aplicación está controlado mediante una lista blanca (*white list*) de IPs, garantizando así el único acceso por parte de las personas colaboradoras dentro del hospital.

En esta subsección se detalla la funcionalidad básica de la aplicación web desarrollada, incluyendo el resto de pantallas en el Anexo B.

3.3.1. Diseño de la interfaz web

Pantalla inicial

La pantalla inicial contiene una breve descripción del proyecto y el botón ‘Entrar’ para acceder al sistema de validación. Además incluye la descripción de las personas que han colaborado en el proyecto, pero puesto que es una pantalla grande en la que es necesario hacer *scroll*, solamente se muestra en la Figura 3.3 una parte de ésta. El resto se incluye en el Anexo B (Figura B.1).



Figura 3.3: Pantalla inicial del sistema de validación.

Pantalla para validación

Para proceder a visualizar los datos para revisar, el experto tiene la posibilidad de introducir el identificador de un paciente concreto o, por el contrario, el sistema muestra uno aleatoriamente que no haya sido revisado. También se incluyen unas flechas para poder navegar por los distintos

pacientes de la base de datos pendientes de revisar, y se ofrece la posibilidad de modificar un paciente ya validado o rechazado.

En esta pantalla (Figura 3.4) el experto debe revisar que los datos extraídos de la Historia Clínica, mostrada en la parte izquierda, se corresponden con los datos de los 46 atributos, mostrados a la derecha. Además, dispone de un campo para comentarios e introducir cualquier aspecto a mejorar del sistema en cada caso.

Por último, contiene un botón para ‘Validar’ o ‘Rechazar’ el contenido del paciente. En caso de validarlo, la aplicación se conecta directamente al sistema basado en reglas (descrito en la Sección 4) y emite el resultado propuesto por este modelo (véase Figura B.3).

The screenshot displays the 'Sistema de Validación' interface. At the top, a dark navigation bar contains the logo 'INDUCESS' and links for 'Inicio', 'Sistema de Validación', 'Estadísticas', and 'Equipo'. Below this, a search bar with the text 'Buscar otro paciente' and a blue 'Buscar' button is visible. A status indicator shows 'Revisados: 101/10487' and a blue arrow icon labeled 'Siguiente'. The main heading is 'Identificador del paciente a revisar' followed by the ID 'f392c479-eb14-3ed0-9664-2b8cce401770'. The interface is divided into two columns:

- Historia Clínica:** Contains a scrollable text area with medical history data, including sections like 'DONACIÓN DE LA SANGRE DEL CORDÓN UMBILICAL:' and various clinical observations.
- Formulario Clínico:** Contains a form with fields for:
 - Edad gestacional (semanas): 39.8571
 - Edad materna (años): 34
 - Raza: Caucásica
 - IMC al inicio de la gestación: 19.433
 - Fumadora: Radio buttons for 'Si' and 'No' (selected).

Figura 3.4: Pantalla para validación del sistema.

Entre otras opciones disponibles en la aplicación se pueden consultar algunos datos estadísticos, información detallada del proyecto así como de las entidades colaboradoras o enviar un correo electrónico de consulta.

4

Sistema basado en reglas

En la actualidad no existen herramientas que apoyen al médico experto en la toma de decisiones dentro del campo de la Obstetricia. A fin de mejorar esta situación se tiene como objetivo construir un sistema de ayuda a la toma de decisiones basada en reglas de decisión provenientes de un panel de expertos en el campo. De manera que, partiendo de la información clínica de los pacientes previamente categorizada en variables, se elaborará un sistema que contenga el conocimiento representado mediante variables y reglas que definen el problema, todo ello obtenido únicamente del propio conocimiento clínico y experiencias.

En este capítulo se describen ciertos componentes que conforman el sistema experto basado en reglas. En primer lugar, se recopilan y analizan las reglas proporcionadas por los expertos y se transforman a un lenguaje entendible por el sistema. Además se describe el método utilizado para el tratamiento de prioridades y el proceso de validación y verificación de las reglas. Finalmente, se diseña y se describe la arquitectura del sistema construido.

4.1. Análisis y diseño de reglas

Uno de los puntos más importantes y críticos en la implementación de la herramienta es transformar el conocimiento del experto en un lenguaje entendible por el sistema, ya que el conocimiento introducido es el que realmente se maneja y a partir del cual se emite el diagnóstico. Por este motivo antes de construir el sistema experto es necesario comprender el dominio y familiarizarse con el tema que se va a tratar. Para ello, no sólo se cuenta con bibliografía especializada en el campo, sino que se extrae conocimiento mediante conversaciones con el experto, donde se formulan reglas en base a su criterio en la práctica clínica y su propia experiencia.

El conocimiento humano no es sencillo de abstraer y mucho menos de representar. Este hecho unido a que en ocasiones la información está incompleta, puede llevar al sistema a no ser capaz de tratar estas circunstancias, ya que sigue una estrategia procedimental. Por ello es necesario realimentar las reglas de acuerdo a la experiencia y comprobar que la base de conocimiento es completa. Para este fin se realizan reuniones iterativas con los expertos donde se verifica que el conjunto de reglas elaborado para ser introducido en el sistema es correcto y coincide con lo expresado por ellos.

Para la implementación de este sistema se ha utilizado la herramienta CLIPS, ya que entre sus características destaca por proporcionar una estrategia de inferencia de encadenamiento hacia adelante, es decir, comienza con una evidencia inicial y se dirige hacia la obtención de la solución, lo cual permite simular el razonamiento deductivo habitual del experto. Para ello, se utilizan reglas de inferencia del tipo *Modus Ponens* ('Si evidencia E entonces hipótesis H '), que permiten obtener conclusiones simples. En este tipo de reglas se examina la evidencia y, si es cierta, la hipótesis pasa a formar parte del conocimiento.

La base de conocimiento que define el problema, inicialmente, se compone de un total de 41 reglas definidas en función de las variables descritas en el apartado 3.1.2 y fueron proporcionadas en lenguaje natural por el experto en Obstetricia. El proceso a seguir por tanto es la introducción de estas reglas en la herramienta manteniendo un *feedback* constante por parte del experto.

4.1.1. Tratamiento de prioridades

El conocimiento que manejan los expertos en el problema del diagnóstico puede ser impreciso e, incluso, presentar conflictos, ya que dependiendo de cada paciente la regla de inferencia empleada para el diagnóstico puede ser diferente. Unido a esto, en el espacio de atributos propuesto no todos son igual de relevantes a la hora de tomar la decisión de iniciar la inducción al parto.

Por ese motivo, por un lado, además de distinguir dos categorías tal y como se ha explicado en 3.1.2, es necesario establecer pesos que indiquen la relevancia de cada atributo predictor. Estos pesos también son proporcionados por el personal experto en el campo y se muestran en la Tabla 4.1.

Tabla 4.1: Distribución por pesos de los atributos predictores del modelo. Los atributos de resultado no tienen ningún peso.

	Atributos
Peso 5-6	<i>motivo_induccion</i> <i>rpm_ingreso</i> <i>bishop_ingreso</i> <i>peso_fetal_estimado</i> <i>crecimiento_fetal</i> <i>edad_gestacional</i> <i>cesareas_previas</i>
Peso 3-4	<i>partos_vaginales_previos</i> <i>numero_fetos</i> <i>motivo_cesarea_previa</i>
Peso 1-2	<i>imc_inicio</i> <i>liquido_amniotico</i> <i>dbp</i> <i>talla_materna</i> <i>tratamiento_fertilidad</i> <i>edad_materna</i> <i>cervicometria</i> <i>complicaciones_embarazo</i>
Peso 0	<i>raza</i> <i>tabaquismo</i>

Por otro lado, el motor de inferencia del sistema experto sigue un mecanismo de razonamiento específico y, dependiendo del orden de aplicación de las reglas, el diagnóstico puede ser diferente. Es por ello necesario utilizar una estrategia de resolución de conflictos basada en prioridades. De esta forma el motor de inferencia ordena las reglas según esas prioridades y selecciona una de ellas como la primera, que será la que se ejecute en el siguiente ciclo de control.

Para la asignación de dicha prioridad, puesto que se dispone de los pesos de cada una de las variables que conforman las reglas, se utiliza la misma idea de los *factores de certeza* utilizada en el sistema MYCIN, explicado en detalle en la Sección 2.3.1.

En este caso el factor de certeza de la evidencia E se tomará como el peso (*weight*) del atributo A que compone la evidencia de la regla. Gracias a la propagación, en este caso, de los pesos de los atributos se determina la prioridad o factor de certeza de la conclusión de cada regla, del siguiente modo: se toma el opuesto del peso del atributo en caso de que éste aparezca negado, el mínimo de los pesos normalizados para la conjunción y el máximo para la disyunción.

Se debe recordar que en el sistema MYCIN el factor de certeza CF tomaba valores entre -1 y 1. En este proyecto se tienen pesos entre 0 y 6, por lo que es necesario normalizarlos entre esos valores, del siguiente modo:

$$CF(\neg A) = -\frac{weight(A) - \max[weights]/2}{\max[weights]/2} \quad (4.1)$$

$$CF(A_1 \wedge A_2) = \frac{\min[weight(A_1), weight(A_2)] - \max[weights]/2}{\max[weights]/2} \quad (4.2)$$

$$CF(A_1 \vee A_2) = \frac{\max[weight(A_1), weight(A_2)] - \max[weights]/2}{\max[weights]/2} \quad (4.3)$$

Gracias a la asignación de prioridades a las conclusiones y, por consiguiente, prioridades de las reglas, es más fácil asemejar la inferencia que realiza el sistema con el experto, teniendo en cuenta que nunca lo podrá sustituir.

4.1.2. Verificación de reglas

En ciertas situaciones complejas, incluso el personal experto puede dar información inconsistente. Por ello, es muy importante controlar la coherencia del conocimiento tanto durante la construcción de la base de conocimiento como durante los procesos de adquisición de datos y razonamiento. Si la base de conocimiento contiene información inconsistente, es muy probable que el sistema experto se comporte de forma poco satisfactoria y obtenga conclusiones absurdas.

El principal objetivo de la verificación es asegurar que el sistema proporcione la respuesta correcta cuando se resuelve un problema dado, lo que involucra habilidad, seguridad y usabilidad a la hora de utilizar dicha herramienta. Para este fin, se deben detectar y, por tanto, eliminar ciertos errores o insuficiencias del sistema como son errores en la semántica durante la implementación de la base de conocimiento, redundancia de reglas, conflicto entre ellas o incluso aquellas innecesarias. De esta forma se asegura un sistema consistente y completo.

Para verificar la coherencia del sistema se ha construido una ontología OWL utilizando Protégé con el objetivo de comprobar los siguientes puntos:

- Redundancia de reglas
- Conflicto de reglas
- Reglas innecesarias

A continuación se procede a describir en detalle cada uno de los puntos anteriores.

Redundancia de reglas

Dos reglas se consideran sintácticamente redundantes si tienen evidencias E y concluyen con hipótesis H idénticas o, su sintaxis no es exactamente igual pero su significado es el mismo.

En el conjunto de reglas se detectaron dos que eran redundantes, pues ya existían otras reglas que se formulaban de diferente manera pero su significado era exactamente el mismo.

Conflicto de reglas

Un conflicto entre reglas ocurre cuando las evidencias E de las reglas son equivalentes pero conducen a conclusiones H distintas.

Gracias a la ontología generada se han podido detectar dos parejas de reglas que producían conclusiones contradictorias. El problema en este caso es que ambas reglas tienen la misma prioridad asignada y una de ellas engloba a la otra, de manera que, el motor de inferencia no tiene ningún criterio para decidir cuál es la que produce el resultado correcto. Para solucionar estos dos casos, y tras confirmar con el experto, se ha modificado la prioridad de la regla más general con un valor más bajo, dentro del rango en que se mueve el factor de certeza, es decir, se ha añadido -1 .

Reglas innecesarias

En las revisiones del conjunto de reglas con el experto en el campo, además se detectaron cuatro reglas que eran innecesarias en la base de conocimiento, bien porque nunca se iban a alcanzar o bien porque no eran correctas. Éstas se han eliminado del conjunto.

Tras el proceso de verificación expuesto, se ha reducido el número de reglas con las que el sistema infiere diagnósticos, pero de esta forma es posible construir un sistema más consistente y completo.

4.2. Diseño del sistema experto

Para desarrollar la herramienta de ayuda a la toma de decisiones en Obstetricia ha sido necesario diseñar un sistema experto, donde la información de los pacientes forme parte de la base de hechos o memoria de trabajo y, el motor de inferencia la utilice para consultar en la base de conocimiento y mostrar el resultado del diagnóstico mediante una interfaz. El esquema del sistema final se puede resumir en la Figura 4.1.

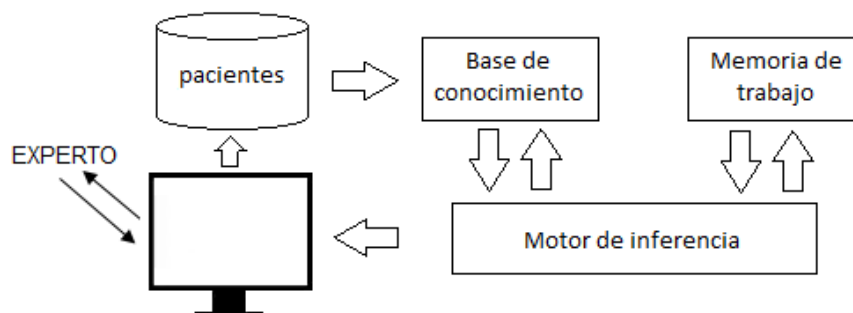


Figura 4.1: Estructura general del sistema experto diseñado.

El programa principal que gestiona el sistema experto ha sido diseñado en el lenguaje de programación Ruby, el cual se conecta a la base de datos para la extracción de información de los pacientes y a la herramienta CLIPS para inferir los resultados.

El funcionamiento de este sistema es el siguiente. El experto, o cualquier otro usuario, desde la interfaz web desarrollada, explicada en la Sección 3.3.1, selecciona el paciente que desee revisar y, en caso de que lo valide, la propia aplicación se encarga de conectar con el sistema experto, el cual emite un resultado y se muestra por pantalla (véase Figura B.3).

Cada uno de los componentes que forman el sistema experto se describe a continuación.

Base de conocimiento

La base de conocimiento contiene, como su propio nombre indica, el conocimiento proporcionado por el experto, codificado mediante las reglas de decisión descritas anteriormente.

Esta base de conocimiento finalmente se compone de 35 reglas simples, las cuales se presentan en el Anexo C y están categorizadas de acuerdo a las variables ya detalladas.

Memoria de trabajo

En la memoria de trabajo, también denominada base de hechos, se almacenan los hechos sobre el problema. Es decir, se modela toda la información extraída de la Historia Clínica Electrónica mediante variables que actúan de entrada al motor de inferencia, éste realiza el respectivo análisis devolviendo otra variable, donde se almacena el resultado del proceso de inferencia.

Motor de inferencia

El motor de inferencia permite controlar el proceso de razonamiento, utiliza datos de la memoria de trabajo y recorre la base de conocimiento haciendo inferencias que le permitan decidir las reglas que satisfacen los hechos. Para ello, prioriza las reglas en función de las prioridades que les han sido asignadas y ejecuta primero aquella con la más alta.

Este motor ha sido configurado con la estrategia de inferencia de encadenamiento hacia adelante, de modo que las variables de entrada determinan el resultado sugerido por el sistema experto.

Además, el motor de inferencia utiliza la información almacenada tanto en la memoria de trabajo como en la base de conocimiento para obtener nuevas conclusiones o hechos y, con ello, incrementar los datos iniciales, ya que incorpora a éstos las nuevas conclusiones obtenidas.

La Figura 4.2 detalla el proceso que realiza el mecanismo de inferencia. Partiendo del programa que gestiona la interfaz web de la aplicación, donde el usuario (bien el médico experto o

cualquier otra persona) indica el paciente del que desea obtener la recomendación y se envía su identificador al programa que gestiona el sistema experto (*main.rb*). Desde éste se invoca a los archivos *patient.clp* y *rules.clp*, los cuales, respectivamente, contienen la plantilla del paciente para almacenar en la memoria de trabajo, y las reglas de inferencia que se almacenan en la base de conocimiento. Este programa se comunica directamente con CLIPS para procesar y emitir una respuesta. Finalmente, la respuesta se envía a la aplicación web, la cual se encarga de mostrar el resultado al usuario en forma de recomendación.

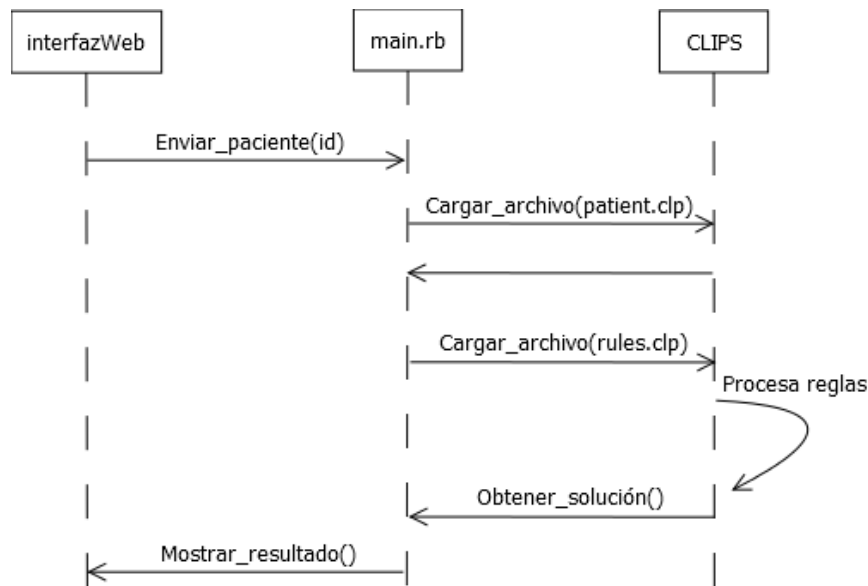


Figura 4.2: Mecanismo de inferencia del sistema experto.

Para finalizar con el diseño del sistema experto, se lleva a cabo un proceso de verificación de su eficacia, considerando los siguientes ámbitos:

- Verificación de la adecuación de la base de conocimiento, comprobando que las reglas que la componen funcionan adecuadamente. Además se comprueba que el sistema experto funciona con razonamiento exacto, es decir, que las variables de entrada son precisas y no existen variables difusas que puedan generar un conflicto.
- Verificación del correcto funcionamiento de la interfaz con el usuario, de modo que exista fluidez en la interacción entre el sistema experto y la base de datos.
- Verificación del funcionamiento del mecanismo de inferencia. Para ello se han creado un total de 35 pacientes ficticios, cada uno de ellos representando una de las posibles combinaciones que es capaz de evaluar el sistema experto. De este modo se asegura la evaluación de todos los casos posibles, comprobando que el sistema genera recomendaciones de forma correcta para esos pacientes.

Por otro lado, la validación del sistema experto se llevará a cabo más adelante, en el apartado de experimentos (Sección 6), donde se mide el acierto entre el sistema y el experto humano,

además de comparar este sistema con los modelos de inferencia basados en algoritmos y técnicas de aprendizaje automático y Big Data.

5

Big Data en medicina

La medicina es uno de los campos del conocimiento que más podrían beneficiarse de una interacción con la computación, mediante la cual se optimizarían ciertos procesos complejos como el diagnóstico clínico. De esto se ocupa el aprendizaje automático junto con el Big Data, que construyen y estudian sistemas capaces de aprender a partir de grandes conjuntos de datos y de mejorar procesos de clasificación y predicción. Para que esos datos se conviertan en conocimiento, necesitan ser procesados y analizados a través de los modelos, pero cada uno de ellos tiene sus propias características, por lo que no todos son adecuados para resolver cualquier tipo de problema.

Concretamente, en el ámbito médico la toma de decisiones es un punto crítico, ya que una decisión errónea influye directamente en la salud de las personas. Por ello, a la hora de utilizar un modelo automático para apoyar los diagnósticos médicos deben estudiarse las ventajas e inconvenientes, así como los riesgos que pueden conllevar. Sin embargo, ningún modelo es mejor que el resto en todas las circunstancias posibles, sino que cada uno tiene sus propias ventajas e inconvenientes dependiendo del problema al que se aplica. No obstante, es muy difícil conocer a priori el modelo más adecuado para un caso determinado. Por ello, se recomienda y, en los experimentos realizados se tendrá en cuenta, probar siempre varios modelos de diferentes paradigmas y así comparar resultados.

En esta sección se presentan algunos modelos de gran difusión en el aprendizaje automático y que se aplican exitosamente al dominio médico, explicando brevemente su funcionamiento y sus características principales. Además se describen las medidas utilizadas en los algoritmos de selección de atributos elegidos para determinar aquellas variables de mayor relevancia en el modelo y, con ello, obtener los mejores resultados posibles.

5.1. Modelos utilizados

Los métodos que se utilicen en el proyecto deben ofrecer ciertas ventajas a la práctica médica que favorezca su uso por parte de expertos. En [28] se revisan las técnicas y modelos de aprendizaje automático que se han aplicado tradicionalmente en medicina y se recogen los requerimientos que deben cumplir para tener éxito en este campo. En la medida de lo posible, se pretende ofrecer una explicación o información adicional que justifique una decisión, ya que puede ayudar de alguna forma comprensible para el especialista en salud a adquirir conocimiento sobre el problema dado.

La mayoría de los modelos utilizados pertenecen al paradigma de aprendizaje supervisado, es decir, necesitan una etapa de aprendizaje para construir un modelo a partir de los datos de entrenamiento etiquetados y después utilizarlo en la predicción sobre instancias desconocidas. Sin embargo, también se estudian otros métodos no supervisados como modelos de *clustering*. A continuación se describen brevemente los siete modelos utilizados, donde los cinco primeros pertenecen al aprendizaje supervisado.

5.1.1. Naive Bayes

Naive Bayes [29] (ingenuo) es un modelo probabilístico fundamentado en el Teorema de Bayes para extraer la regla de clasificación y se basa en la suposición de que todos los atributos son condicionalmente independientes dada la clase. Pero esa suposición se aleja demasiado de la realidad en la mayoría de los problemas, aunque su rendimiento es sorprendentemente bueno debido a que no hay que estimar distribuciones de probabilidad conjuntas. Así, cuando se introduce una nueva instancia en el modelo, la probabilidad asociada a cada clase es la multiplicación de la probabilidad para la correspondiente clase de cada uno de los atributos que conforman la instancia. La clase final asignada es la que más alta probabilidad asociada tiene.

Los clasificadores Bayesianos son uno de los modelos preferidos en el ámbito médico ya que ofrecen gran capacidad para explicar sus predicciones. Determinan en qué medida los valores de cada atributo contribuyen positiva o negativamente a la decisión tomada. Además presentan un rendimiento bastante bueno, manejan los valores perdidos como si fuesen aleatorios, toleran niveles aceptables de ruido y tienen un buen nivel de transparencia (aunque no tanto como los árboles de decisión).

5.1.2. Árboles de decisión

Los árboles de decisión se ubican dentro de una rama del aprendizaje automático denominada aprendizaje simbólico. Esto es, dado un conjunto de datos se elaboran diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la

resolución de un problema. Un árbol de decisión consta de nodos, ramas y hojas, que son como los nodos pero no surge ninguna rama de ellas. De cada nodo nacen dos o más ramas en las que se formulan proposiciones lógicas excluyentes y particionan el espacio de atributos. Esto quiere decir que cada instancia, dependiendo de los valores de los atributos implicados en la proposición lógica de un nodo, cumple la condición correspondiente a una de las ramas. De este modo, comenzando en el nodo raíz, cada instancia se dirige hasta una de las hojas. A todas las instancias pertenecientes a la misma hoja se les asigna una etiqueta común.

Previamente a construir dicho árbol, debe existir una fase de entrenamiento en la que se deciden las preguntas que se van a formular en cada nodo. Para ello, existen diversos algoritmos, entre ellos, *C4.5* que será el utilizado en los experimentos de este proyecto.

El algoritmo *C4.5* es una extensión de *ID3* que permite el uso de atributos numéricos continuos, tiene en cuenta los valores desconocidos y realiza un proceso de poda inteligente del árbol, es decir, limita o recorta la profundidad del árbol para reducir su tamaño y permitir mayor capacidad de generalización.

En general, los árboles de decisión presentan un rendimiento apropiado, así como un manejo aceptable de valores perdidos y ruido. También reducen la cantidad de datos necesaria para evaluar a un paciente, ya que no suelen aparecer todos los atributos de entrenamiento en el árbol construido. Sin embargo, las claves del éxito de los árboles de decisión son su gran transparencia (se puede observar toda la estructura del conocimiento extraído de los datos) y su capacidad de explicar cada una de las predicciones (registrando cómo la instancia recorre el árbol hasta llegar a la hoja correspondiente). Adicionalmente, en producción no es necesario realizar ningún cálculo y se pueden utilizar sin el ordenador, lo que es muy bien recibido por los profesionales médicos no familiarizados con la informática.

5.1.3. Redes neuronales

Las redes neuronales se han vuelto populares en medicina debido a su flexibilidad y dinamismo. Inicialmente, surgieron por la motivación de tratar de reproducir el funcionamiento del cerebro.

El modelo de red neuronal *Perceptrón Multicapa* (*Multilayer Perceptron*) incluye varias capas de neuronas artificiales interconectadas entre sí. La etapa de entrenamiento consiste en, dado un conjunto de pesos iniciales, introducir cada uno de las instancias de entrenamiento en el modelo y comparar el valor de salida con la clase real esperada. Dependiendo del error de la clase inferida, el algoritmo *Backpropagation* ajusta los pesos de la red propagando el error, desde la capa de salida a la de entrada, para hacer que la salida inferida sea igual a la esperada. Esto es conocido como *descenso por gradiente* del error de predicción.

Aunque el *Multilayer Perceptron* no ha alcanzado una precisión perfecta ofrece un rendimiento razonablemente bueno. Los principales inconvenientes para su aplicación en el ámbito

médico son su falta de transparencia y su escasa capacidad de explicar las predicciones emitidas. En la Sección 2.4 del estado del arte se han expuesto algunas técnicas que permiten explicar las predicciones, incluso considerando el modelo como una *caja negra*. De esta manera, se puede mitigar el impacto negativo de los puntos débiles de las redes neuronales y con ello aprovechar mejor su excelente rendimiento.

5.1.4. Máquinas de vectores soporte

Las *Máquinas de Vectores Soporte (SVM)* tienen como objetivo ajustar un conjunto de parámetros, que permiten establecer fronteras en el espacio de n dimensiones y aproximar funciones o separar patrones en diferentes regiones del espacio de atributos. En este caso, el método de entrenamiento para ajustar los parámetros es la maximización del margen existente entre el hiperplano separador y las instancias de las dos clases (inicialmente, este modelo se diseñó para resolver problemas de clasificación de dos clases pero hay extensiones para problemas multiclase).

De manera similar a las redes neuronales, las *SVM* presentan un buen rendimiento pero el hecho de que el entrenamiento consiste en el ajuste de parámetros ocultos y la predicción únicamente en un resultado numérico, hace que la transparencia y la capacidad de explicación sean pobres. Este aspecto se mejora mediante la aplicación de las técnicas ya explicadas en la Sección 2.4 del estado del arte.

5.1.5. Random Forest

Se trata de una combinación de árboles de decisión aleatorios en la que cada árbol utiliza un subconjunto de datos diferentes de entrenamiento en cada iteración. El *Random Forest* se basa en la idea esencial del *bagging* de promediar muchos modelos ruidosos pero aproximadamente imparciales y, por tanto, reducir la posible variación.

El funcionamiento de este modelo es el siguiente. En cada nodo de los árboles de decisión construidos, se elige la mejor subdivisión hecha por un subconjunto de variables explicativas seleccionadas aleatoriamente. Los árboles que se obtienen son maximales, es decir, que en este caso no se podan. De manera que, la predicción de una observación emitida con este modelo es la clase más votada entre las predicciones hechas por los árboles.

Random Forest se considera uno de los algoritmos con mejor rendimiento especialmente en problemas que cuentan con una cantidad importante de variables explicativas [30]. Tiene un método eficaz de valores perdidos y ruido que le permite mantener la exactitud cuando una gran proporción de datos son desconocidos. Además aunque es capaz de proporcionar las variables más importantes en la clasificación, a diferencia de los árboles de decisión, su salida es difícil de interpretar.

5.1.6. Vecinos más cercanos

Es un modelo perteneciente al paradigma de aprendizaje perezoso, que se caracteriza por la ausencia de una fase de extracción de conocimiento de los datos. El funcionamiento está dirigido por la demanda de clasificación de una nueva instancia. En ese momento se analiza el entorno de dicha instancia y se emite una predicción basada, como el propio nombre del algoritmo indica, en las instancias vecinas más cercanas a la demandada.

El algoritmo de predicción para un nuevo paciente, desconocido hasta el momento, es muy sencillo. En primer lugar, se fija un natural k que determina el número de vecinos que se tendrán en cuenta y se hallan las k instancias más cercanas a la dada, empleando alguna distancia (ya sea Euclídea, Mahalanobis, etc.). A continuación, la clase mayoritaria de entre las k instancias más similares a la de entrada es la clase inferida para la misma.

En general, k -vecinos más cercanos (conocido como kNN) maneja aceptablemente los valores perdidos y el ruido pero no presenta un rendimiento claramente mejor que ninguno de los modelos anteriores. Su transparencia es pobre ya que, por ser un algoritmo de aprendizaje perezoso, se limita únicamente a emitir una predicción como respuesta a una demanda. El factor diferenciador que hace interesante el uso de este modelo es la forma de explicar las predicciones. Una nueva instancia se etiqueta en función de cómo están etiquetadas sus vecinas. Haciendo una analogía, se puede identificar esta aproximación con el método actual de diagnóstico clínico basado en la experiencia adquirida al tratar casos similares (vecinos más cercanos).

5.1.7. Maximización de la esperanza

El algoritmo de maximización de la esperanza (*Expectation-Maximization*, EM) pertenece a una familia de modelos que se conocen como *Finite Mixture Models*, los cuales se pueden utilizar para segmentar conjuntos de datos. Dentro de la clasificación destaca frecuentemente en algoritmos de *clustering*.

EM es una técnica utilizada para encontrar el estimador de máxima verosimilitud de una distribución de un conjunto de datos incompletos. Para ello, alterna pasos de esperanza (paso E), donde se obtiene la esperanza de la verosimilitud mediante la inclusión de variables latentes como si fuesen observables, y un paso de maximización (paso M), donde se calculan los estimadores de máxima verosimilitud de los parámetros mediante la maximización de la verosimilitud esperada del paso E anterior. Los parámetros que se obtienen de este último paso M son utilizados para el paso E siguiente, y así el proceso se repite hasta converger. Finalmente se obtiene un conjunto de *clusters* que agrupan el conjunto de datos original.

El éxito de este algoritmo depende en gran medida de los valores iniciales del conjunto de parámetros, aunque tiene gran capacidad para manejar información desconocida y determinar variables ocultas.

5.2. Selección de atributos

Entre los objetivos de este proyecto destacaba valorar la efectividad de las variables y determinar aquellas de mayor relevancia en el modelo. En general, suele ocurrir que existan datos irrelevantes o redundantes. Si no se detecta y se elimina esa información con menor valor predictivo antes de entrenar un modelo de aprendizaje automático, su rendimiento puede verse mermado; bien reduciendo la precisión de la predicción o, bien aumentando el coste computacional (operar con más información de la necesaria puede conllevar a obtener un modelo demasiado complejo y se produzca sobreajuste). Es por ello que antes de utilizar los modelos predictivos se reduce la dimensionalidad del problema, a fin de seleccionar las variables más importantes además de facilitar el ajuste de un modelo y aportar con ello un valor añadido a la investigación biomédica. De este modo se pueden descubrir factores asociados a determinadas enfermedades o dependencias desconocidas entre diversos atributos.

En este subapartado se describen, en primer lugar, las medidas de evaluación de atributos empleadas en el proyecto y, a continuación, los algoritmos de selección de acuerdo a esa medida.

5.2.1. Medidas de evaluación

Se debe tener en cuenta que el conjunto óptimo de atributos es siempre relativo a un criterio de evaluación, es decir, un subconjunto óptimo elegido según una medida de evaluación puede cambiar al utilizar otra medida distinta. Sin embargo, en la práctica se comprueba con cierta frecuencia, que si un atributo es relevante aparece en varios subconjuntos escogidos según distintas funciones de evaluación.

- Medidas de distancia

Estas medidas, también conocidas como medidas de separabilidad, divergencia o discriminación, estiman la capacidad que posee un subconjunto de atributos para separar las clases. El conjunto de atributos que seleccionan es aquel que minimiza la distancia intra-clase y maximizan la inter-clase. Existen muchas medidas de distancia, éstas pueden ser Euclídea, Mahalanobis o Kolmogorov, entre otras.

- Medidas de información

Este tipo de medidas están basadas en la ganancia de información de un atributo, esto es, la entropía de la teoría de la información, que es una medida de la incertidumbre de una variable aleatoria. Entre las medidas de información más frecuentes se encuentran la entropía de Shannon, de Renyi, cuadrática o información mutua.

- Medidas de dependencia

Las medidas de dependencia o correlación evalúan la capacidad de predecir el valor de una determinada variable en función de otra. El coeficiente de correlación es una medida de dependencia clásica utilizada para calcular la correlación entre un atributo y la clase, seleccionando aquellos atributos con mayor correlación.

5.2.2. Algoritmos de selección

En función de las medidas de evaluación descritas anteriormente, se presentan en este apartado los cinco algoritmos utilizados para la selección de atributos. Todos ellos son de tipo *filter*, es decir, no requieren el entrenamiento de ningún modelo para decidir cuáles son los atributos más importantes. Muchos de estos algoritmos no tienen un nombre especial por el que se les reconoce sino que lo toman de la búsqueda realizada, en estos casos, se hará referencia a ellos por el nombre utilizado en la herramienta Weka [31], empleada en los experimentos.

- ReliefF

Relief es un algoritmo inspirado en el aprendizaje basado en casos que obtiene los atributos estadísticamente más relevantes. Su funcionamiento se basa en asignar un peso a cada atributo y seleccionar aquellos que superen un umbral prefijado, esos serán los atributos que mejor permitan discernir entre las clases a las que pertenece cada instancia.

Para este propósito, en cada instancia aleatoria, analiza el espacio de atributos localmente y busca dos vecinos más cercanos (de la misma clase y de distinta clase). El peso asociado a un atributo se modifica a partir de la distancia euclídea entre el valor del atributo de la instancia y el valor del mismo atributo de los vecinos encontrados.

Pero este algoritmo presenta ciertas limitaciones que, posteriormente, se superan mediante varias extensiones propuestas por Kononenko [32]. En sus experimentos comprueba que la que mejor funciona es *ReliefF*, en la que se seleccionan las k instancias más parecidas (de la misma y de diferente clase, respectivamente) y se toma su media. Seleccionar k vecinos en lugar de uno, aporta mayor robustez al método. Sin embargo, no funciona mejor cuanto mayor sea k , ya que si se consideran demasiados vecinos, se pierde la perspectiva local del método y dejan de detectarse ciertas dependencias entre atributos.

- mRMR

mRMR (*Minimum Redundancy Maximum Relevance*) [33] trata de obtener un subconjunto de atributos de tal manera que sean lo más diferente posible entre sí (mínima redundancia) y que aporten la máxima información posible sobre la variable objetivo (máxima relevancia).

Es un algoritmo iterativo que determina la relevancia de un atributo a partir de la entropía con su clase y la redundancia del mismo, calculando la entropía entre las variables

seleccionadas y las que todavía no se han escogido. Como criterio de selección se utiliza la diferencia de estos dos valores, escogiendo aquellos que tengan valores positivos más altos (ya que interesa que el valor de relevancia sea mayor que el de redundancia).

- Razón de ganancia

El algoritmo *GainRatioAttributeEval* [34] evalúa el peso de cada atributo midiendo su razón de beneficio (*gain ratio*) con respecto a la clase. Esta medida tiene en cuenta tanto la ganancia de información como las probabilidades de los distintos valores de cada atributo, recogidas como la entropía del conjunto de datos respecto a los valores del atributo.

En base a ese criterio, este algoritmo premia a aquellos atributos que, teniendo una ganancia de información menor disponen también de un número menor de valores.

- Ganancia de información

El algoritmo *InfoGainAttributeEval* [34] consiste en la evaluación de los atributos mediante la ganancia de información de cada uno con respecto a la clase, discretizando antes aquellos atributos que son numéricos.

- Correlación

El algoritmo *CorrelationAttributeEval* [34] evalúa el valor de un atributo mediante la medida de correlación de Pearson entre el mismo y la clase.

6

Evaluación y análisis de resultados

Con el objetivo de evaluar los sistemas desarrollados y determinar aquellas variables relevantes en el proceso de inducción, se ha llevado a cabo una etapa experimental con su correspondiente análisis de resultados.

En primer lugar, se presentan los resultados del proceso de validación llevado a cabo por los expertos del Hospital de Fuenlabrada sobre el sistema de recuperación de información implementado. En segundo lugar, se presentan diferentes conjuntos de datos contruidos tras la selección de atributos, los cuales son utilizados para evaluar los modelos de aprendizaje automático. Para ello, se describen las métricas empleadas así como los experimentos realizados, incluyendo una comparativa entre los dos sistemas propuestos. A continuación, se analizan detalladamente algunas de las técnicas de aprendizaje automático utilizadas y se muestran los resultados obtenidos en forma de gráficas o tablas. Por último, se realiza un análisis de dichos resultados, a fin de mostrar las ventajas que pueden ofrecer tanto a expertos en modelos como a cualquier usuario que utilice un sistema de aprendizaje automático como apoyo a su labor.

Para realizar las diferentes pruebas se ha empleado el sistema operativo Linux. Todo el código utilizado se lanza mediante scripts de Bash, Ruby y Python [35], que en algunas ocasiones actúan como meros intermediarios para invocar a otros programas pero, en otras, ejecutan parte de los algoritmos y técnicas desarrolladas. También se utilizan otras herramientas como Awk y R [36] para la generación de explicaciones y visualización de resultados.

6.1. Preparación de datos

Como ya se introdujo en la Sección 3.1.2, el espacio de atributos se puede separar en dos categorías en función de si los atributos son predictores o, por el contrario, de resultado en la decisión de inducción. En base a esta separación surge la posibilidad de diseñar diferentes modelos de datos para representar un paciente, esto es, escoger distintos atributos que permitan evaluar el conjunto que mejor representa el problema estudiado.

Para ello se han diseñado tres modelos diferentes de paciente, los cuales se detallan a continuación.

Modelo 1 (Falsos predictores)

- **Objetivo:** Representar la información completa recopilada de la Historia Clínica Electrónica, teniendo en cuenta que entre sus atributos se encuentran los denominados de resultado.

Es decir, contiene tanto atributos predictores como aquellos que aparecen durante o después del parto y son falsos predictores, los cuales no aportan ninguna información de clasificación. Más bien, contiene atributos altamente correlacionados con la clase, por lo que se espera que el error de predicción sea muy cercano a cero. Este modelo no es representativo para un problema real pues se predice sabiendo lo que ha pasado.

- **Atributos:** Se incluyen en la Tabla 6.1.

Tabla 6.1: Atributos que componen el Modelo 1 (Falsos predictores).

<i>id</i>	<i>dbp</i>	<i>hipertonias</i>
<i>edad_gestacional</i>	<i>motivo_induccion</i>	<i>epidural</i>
<i>edad_materna</i>	<i>peso_fetal_estimado</i>	<i>tipo_parto</i>
<i>raza</i>	<i>bishop_ingreso</i>	<i>motivo_parto_instrumental</i>
<i>imc_inicio</i>	<i>numero_fetos</i>	<i>motivo_cesarea</i>
<i>tabaquismo</i>	<i>induccion_parto</i>	<i>ph_cordon_arterial</i>
<i>talla_materna</i>	<i>tipo_induccion</i>	<i>apgar1</i>
<i>tratamiento_fertilidad</i>	<i>oxitocina_maxima</i>	<i>apgar5</i>
<i>partos_vaginales_previos</i>	<i>oxitocina_horas</i>	<i>peso_neonatos</i>
<i>cesareas_previas</i>	<i>propess_horas</i>	<i>tipo_reanimacion</i>
<i>cervicometria</i>	<i>misofar_dosis</i>	<i>traslado_neonato</i>
<i>liquido_amniotico</i>	<i>tiempo_inicio_parto</i>	<i>dia_parto</i>
<i>rpm_ingreso</i>	<i>tiempo_fin_parto</i>	<i>hora_parto</i>
<i>motivo_cesarea_previa</i>	<i>horas_induccion</i>	<i>fecha_entrada</i>
<i>complicaciones_embarazo</i>	<i>alteraciones_fcf</i>	<i>crecimiento_fetal</i>

Modelo 2 (Atributos predictores)

- **Objetivo:** Generar un modelo que aporte la máxima información simulando un comportamiento real del sistema, ya que a la hora de tomar la decisión de inducción no se conocen las variables de resultado que el modelo anterior consideraba. De este modelo se eliminan ciertas variables como el *id* o las fechas, entre otras.
- **Atributos:** Se incluyen en la Tabla 6.2.

Tabla 6.2: Atributos que componen el Modelo 2 (Atributos predictores).

<i>edad_gestacional</i>	<i>partos_vaginales_previos</i>	<i>crecimiento_fetal</i>
<i>edad_materna</i>	<i>cesareas_previas</i>	<i>dbp</i>
<i>raza</i>	<i>cervicometria</i>	<i>motivo_induccion</i>
<i>imc_inicio</i>	<i>liquido_amniotico</i>	<i>bishop_ingreso</i>
<i>tabaquismo</i>	<i>rpm_ingreso</i>	<i>peso_fetal_estimado</i>
<i>talla_materna</i>	<i>motivo_cesarea_previa</i>	<i>numero_fetos</i>
<i>tratamiento_fertilidad</i>	<i>complicaciones_embarazo</i>	

6.1.1. Selección de atributos

Con el fin de eliminar la información irrelevante que existe en el conjunto de datos y, con ello, reducir la dimensionalidad del problema, se lleva a cabo una selección de atributos. Puesto que se ha visto que el Modelo 1 no es representativo para un problema real, solamente se realiza dicha selección y se eliminan algunas variables del Modelo 2. Las variables resultantes de la ejecución de los algoritmos descritos en la Sección 5.2.2 serán las que conformen otro nuevo modelo (Modelo 3) utilizado en los experimentos.

Los cinco algoritmos de selección de atributos se ejecutan a través de la herramienta Weka con los siguientes parámetros:

- **ReliefF:** 10 vecinos más cercanos.
- **mRMR:** Configuración por defecto de Weka. Cálculo de los pesos por diferencia de información mutua.
- **GainRatioAttributeEval:** Configuración por defecto de Weka.
- **InfoGainAttributeEval:** Configuración por defecto de Weka.
- **CorrelationAttributeEval:** Configuración por defecto de Weka.

Modelo 3 (Atributos predictores tras selección)

- **Objetivo:** Eliminar las variables irrelevantes y construir un modelo únicamente con aquellas más influyentes a la hora de tomar la decisión de realizar una inducción. Este modelo se construye seleccionando las seis variables resultantes de la selección de atributos sobre el Modelo 2. Puesto que se han utilizado cinco algoritmos diferentes, se obtienen cinco submodelos de paciente.
- **Atributos:** En la Tabla 6.3 se muestran los seis atributos más relevantes que se han escogido en función de los resultados de la selección de atributos.

Tabla 6.3: Atributos que componen el Modelo 3 (Atributos predictores tras selección), obtenido seleccionando las variables resultantes de los algoritmos ReliefF, mRMR, *GainRatioAttributeEval*, *InfoGainAttributeEval* y *CorrelationAttributeEval* de selección de atributos sobre el Modelo 2.

ReliefF	mRMR	GainRatioAttributeEval
<i>motivo_induccion</i>	<i>raza</i>	<i>rpm_ingreso</i>
<i>bishop_ingreso</i>	<i>partos_vaginales_previos</i>	<i>motivo_induccion</i>
<i>partos_vaginales_previos</i>	<i>motivo_cesarea_previa</i>	<i>cesareas_previas</i>
<i>raza</i>	<i>motivo_induccion</i>	<i>motivo_cesarea_previa</i>
<i>edad_materna</i>	<i>bishop_ingreso</i>	<i>partos_vaginales_previos</i>
<i>complicaciones_embarazo</i>	<i>numero_fetos</i>	<i>numero_fetos</i>

InfoGainAttributeEval	CorrelationAttributeEval
<i>motivo_induccion</i>	<i>motivo_induccion</i>
<i>rpm_ingreso</i>	<i>rpm_ingreso</i>
<i>bishop_ingreso</i>	<i>bishop_ingreso</i>
<i>edad_gestacional</i>	<i>motivo_cesarea_previa</i>
<i>partos_vaginales_previos</i>	<i>cesareas_previas</i>
<i>motivo_cesarea_previa</i>	<i>partos_vaginales_previos</i>

Con la evaluación de los modelos de paciente propuestos se pretende determinar el conjunto de atributos que mejor predice el éxito de la inducción al parto y, en función de éste, encontrar el algoritmo de aprendizaje automático más preciso, siendo capaz de explicar la clasificación obtenida con el mismo. A partir de esos resultados se realiza una comparación con el modelo de referencia basado en reglas provenientes del experto.

Para llevar a cabo los diferentes experimentos se dispone de la información recopilada mediante el sistema de recuperación de información. Es decir, como ya se ha comentado, 10.487 pacientes, los cuales están categorizados por diferentes variables dependiendo del modelo de paciente utilizado. A modo de resumen, en la Tabla 6.4 se muestran los modelos de paciente utilizados para evaluar los algoritmos. Todos ellos están compuestos por las 10.487 instancias con atributos tanto numéricos como nominales.

Tabla 6.4: Modelos de paciente sobre los que se aplican técnicas de aprendizaje automático. *a* indica el número de atributos (numéricos y nominales). En todos ellos hay un total de 10.487 instancias.

Nombre	a
Modelo 1 (Falsos predictores)	45
Modelo 2 (Atributos predictores)	20
Modelo 3 (Atributos predictores tras selección)	6

6.2. Métricas de evaluación

Para realizar una evaluación exhaustiva se ha diseñado una metodología de evaluación, la cual se detalla en esta sección. Los objetivos que se van a evaluar son: el correcto funcionamiento del sistema de recuperación de información diseñado, los resultados de la inferencia llevada a cabo por los dos sistemas descritos y la explicación de la clasificación obtenida con algunos de los modelos que mejores resultados proporcionan.

Resultados de validación del sistema IR

El correcto funcionamiento del proceso de extracción se ha comprobado con varias pruebas de concepto de manera incremental e iterativa, comenzando de manera *offline* hasta llegar a ser *online*. Para ello, se utiliza la medida de error de validación, es decir, el porcentaje de instancias rechazadas con respecto al total de revisadas.

En una primera revisión *offline* y antes de contar con la presencia del profesional médico, se han validado de manera unitaria cada una de las funcionalidades de la aplicación así como una muestra reducida de pacientes, en torno a 10 instancias. En estas revisiones se detectaron errores y mejoras que se incluyen al *diccionario* y a los algoritmos de recuperación de información.

En la segunda revisión (Fase 1) se cuenta con la ayuda de tres profesionales en el campo. En esta prueba se ha seleccionado un subconjunto de datos pequeño de 25 pacientes y se ha realizado un proceso de validación por parte de los expertos desde el propio hospital de manera *online*. Tras dicha revisión se realimenta de nuevo el *diccionario* y algoritmos del sistema de recuperación de información.

Finalmente, una vez añadidas las mejoras al sistema de extracción de información, se repite la prueba de concepto anterior (Fase 2) y se revisan en torno a 100 pacientes. De igual modo, se realimenta el sistema con el objetivo de obtener una solución lo más exhaustiva posible.

Resultados de inferencia

El objetivo de la herramienta de ayuda a la toma de decisiones es resolver un problema de clasificación real. Consiste en predecir el éxito de un parto mediante *Inducción*, *No inducción* o *Cesárea*. Las clases (variable *resultado*) dentro del conjunto de datos no están balanceadas ya que las proporciones son 55,32 % *No inducción*, 26,31 % *Inducción* y 18,38 % *Cesárea*, y a pesar de no ser una diferencia grande, dificulta la generación de un buen modelo de clasificación como se explicará en el apartado de resultados.

El proceso seguido en la ejecución del sistema basado en técnicas de aprendizaje automático y Big Data se resume a continuación:

- Entrenamiento de los modelos de paciente y técnicas: El conjunto de datos disponible se divide en dos subconjuntos, *train* y *test*. Con los datos de *train* se entrenan diferentes algoritmos que permiten generar el modelo de inferencia.
- Test de los modelos de paciente y técnicas: Utilizando el subconjunto de *test* y los distintos modelos generados en la fase de entrenamiento, se han realizado predicciones del éxito de inducción.
- Evaluación: Para evaluar la calidad de esas predicciones y compararlas con las obtenidas a partir del modelo de referencia (sistema experto basado en reglas), se emplean cinco medidas de error.

Entre estas medidas está el error de clasificación, que compara el valor real conocido de la clase de cada instancia I_i y puede tomar tres valores, es decir, compara $\{\Omega(I_i)\}_{i=1}^n \in \{C_1, C_2, C_3\}, \forall i$, con la clasificación ofrecida por el modelo $\{f(I_i)\}_{i=1}^n \in \{C_1, C_2, C_3\}, \forall i$. Se trata de una medida global y no considera resultados por clase. En problemas con clases desbalanceadas, se utilizan también otras medidas de evaluación basadas en la información suministrada por la matriz de confusión, las cuales se incluyen en el Anexo D. Estos errores son calculados para cada una de las clases y después promediados equitativamente para presentar un único valor en cada caso. Aunque la mayoría de estas medidas se utilizan en problemas de dos clases, en este caso, se utiliza el enfoque multiclase *one vs all*, se crean tantos modelos de clasificación binaria como clases existan en el problema que se está tratando. Dicho enfoque se muestra en las Tablas 6.5 y 6.6.

Tabla 6.5: Esquema de los C problemas de clasificación binaria a realizar.

Espacio	1	2	3
Positiva	C_1	C_2	C_3
Negativa	$C_2 \cup C_3$	$C_1 \cup C_3$	$C_1 \cup C_2$

Tabla 6.6: Representación de la matriz de confusión para cada problema de clasificación binaria.

	Positiva Predicha	Negativa Predicha
Positiva Real	Verdaderos Positivos (TP)	Falsos Negativos (FN)
Negativa Real	Falsos Positivos (FP)	Verdaderos Negativos (TN)

De manera que las métricas utilizadas son:

- Error de clasificación: $ErrClassif = \frac{\#\{\Omega(I_i) \neq f(I_i)\}_{i=1}^n}{n}$.
- Precisión (Valor predictivo positivo): $Prec = \frac{TP}{TP + FP}$, donde TP es el número de instancias bien clasificadas para la clase positiva y $TP + FP$, el número de instancias clasificadas como clase positiva.
- Recall (Ratio de verdaderos positivos): $Recall = \frac{TP}{TP + FN}$, donde $TP + FN$ es el número de instancias de clase positiva.
- Efectividad de la clasificación: $F - measure = \frac{2 \cdot Recall \cdot Prec}{Recall + Prec}$.
- Área bajo la curva ROC: $AUC = \frac{1 + Recall - FPR}{2}$, siendo $FPR = \frac{FP}{FP + TN}$ la fracción de datos mal clasificados en la clase positiva.

La evaluación de los errores se realiza con la técnica de *cross-validation* de 10 *folds* con el fin de garantizar la independencia entre los conjuntos de *train* y test escogidos, y así obtener resultados más estables.

6.3. Resultados de validación del sistema IR

Los resultados de las evaluaciones realizadas por los expertos del hospital se analizan en detalle a continuación.

La validación del sistema se ha llevado a cabo de manera *online*, donde los expertos acceden desde su propio ordenador a la aplicación a través de la web, con el fin de revisar un subconjunto de pacientes seleccionado de manera aleatoria. Como se ha explicado al inicio de esta sección, se realizan dos fases de validaciones, cuyos resultados se presentan en la Tabla 6.7.

Tabla 6.7: Porcentajes de error en el sistema de recuperación de información de cada fase. En la columna etiquetada con n se indica el número de instancias aleatorias del subconjunto a revisar; la siguiente columna corresponde al porcentaje de error en cada fase.

	n	Error de validación (%)
Fase 1	26	34,62
Fase 2	101	29,70

Como se puede observar en la Tabla 6.7 se logra mejorar el proceso de recuperación de información en la segunda fase con respecto al error inicial. Sin embargo, aunque la mejora entre las dos fases no parece muy significativa, repercute directamente en los resultados obtenidos de las reglas de inferencia y los modelos de aprendizaje automático, ya que en cada iteración de la validación se elimina ruido e información irrelevante.

Tras completar la Fase 2, es necesario añadir al *diccionario* los comentarios de los casos rechazados por los expertos. Antes de ello, se han analizado detenidamente cada uno de esos comentarios. Se ha podido observar que un 43 % de los pacientes que fueron rechazados se deben a ciertas variables (de resultado) comunes, las cuales no intervienen en la toma de decisiones o no se han seleccionado como relevantes tras la selección de atributos y, por tanto, no influirán en los resultados de los sistemas implementados. Las variables que destacan en el proceso de validación se incluyen en la Tabla 6.8.

Tabla 6.8: Errores más comunes en las variables extraídas por las que se rechazan pacientes. En la primera columna se indica el atributo y en la segunda el tipo de error a corregir.

Atributo	Tipo de error detectado
<i>tabaquismo</i>	Campo de texto libre. Se escribe de numerosas formas distintas si el paciente fuma, por ejemplo, <i>c/día</i> , <i>cig/día</i> , <i>cigarros/día</i> ...
<i>cervicometria</i>	No se encuentra en los formularios de su Historia Clínica ese valor.
<i>tipo_induccion</i>	No se especifica el tipo de inducción o no contiene el formulario correspondiente en su Historia Clínica.
<i>alteraciones_fcf</i>	Campo de texto libre. Se escribe de muchas formas distintas como son: <i>decelariones</i> , <i>desaceleraciones</i> , <i>bradicardias</i> , <i>no reactivo</i> ...
<i>hipertonias</i>	Campo de texto libre. Aparece de varias formas escrito (<i>hipertonías</i> , <i>taquisistolias</i> , <i>polisistolias</i> ...).
<i>peso_neonatos</i>	No aparece el dato o aparece en otro campo de texto libre.

Teniendo en cuenta que el 43 % de los rechazos se podrían considerar como válidos, bien porque se amplíe el *diccionario* o bien porque no se tiene en cuenta esa variable a la hora de predecir, el sistema podría llegar a mejorar notablemente pasando incluso a un 16,83 % de error. Este hecho supone una gran motivación para continuar con el proceso de validación alguna fase más. Sin embargo, por limitaciones de tiempo en este trabajo no se han considerado más iteraciones, pero se sugiere como trabajo futuro.

Gracias a esta prueba de validación se ha podido comprobar cómo el funcionamiento del sistema es el correcto y el esperado por parte del experto, por lo que se puede concluir que la prueba de concepto final resultó satisfactoria.

6.4. Resultados de inferencia

Tras llevar a cabo un análisis del sistema de recuperación de información y enriquecer los modelos, se presentan en esta subsección los resultados de los experimentos realizados.

En primer lugar, se analizan los resultados obtenidos del sistema experto basado en reglas. A continuación, los obtenidos mediante el sistema basado en algoritmos de aprendizaje automático para, finalmente, realizar una comparativa entre ambos sistemas.

6.4.1. Resultados del sistema experto

Se pretende modelar el error del sistema experto, el cual se basa en reglas de decisión que infieren sobre las variables de los datos recopilados mediante el sistema IR anteriormente validado. Los errores medidos para este modelo considerado de referencia se incluyen en la Tabla 6.9.

Tabla 6.9: Resultados del sistema experto (Modelo de referencia). Las distintas columnas indican las medidas de error del modelo.

Modelo de referencia				
ErrClassif (%)	Prec	Recall	F-measure	AUC
41,89	0,567	0,581	0,571	0,646

Cabe destacar que los errores obtenidos con el sistema basado en reglas pueden estar influenciados por la propia experiencia y conocimientos del experto que las formula, a lo que se suma la dificultad ya explicada de traducir ese conocimiento en lenguaje entendible por el propio sistema. Por ello, tomando como referencia el 41,89 % de error, se pretende encontrar otros modelos que mejoren este resultado y, además, ofrezcan una explicación consistente de la clasificación obtenida, favoreciendo su utilización en el ámbito médico.

6.4.2. Resultados de los modelos de aprendizaje automático

Como se ha explicado al inicio de la sección, los algoritmos empleados, tanto clasificadores como técnicas de *clustering*, se aplican sobre los tres modelos de paciente y los correspondientes submodelos descritos en la subsección 6.1 para predecir la clase. Dichos algoritmos utilizados son los descritos en la Sección 5.1 mediante la herramienta Weka con la siguiente configuración:

- Naive Bayes: Configuración por defecto de Weka.
- C4.5: Configuración por defecto de Weka del clasificador J48 (confidence factor=0,25, minNumObj=2, unpruned=false).
- Multilayer Perceptron: Configuración por defecto de Weka (learning rate=0,3, momentum=0,2, hidden layers=(atributos (a)+clases (3))/2, training time=500, validation threshold=20).
- SVM: clasificador LibSVM (SVM type=C-SVC, degree=3, kernel type=radial basis function: $\exp(-\gamma \cdot |u - v|^2)$, tol=0,001, normalize=false, not replace missing values=false). En algunos casos, γ y C se obtienen mediante *cross-validation* de 10 *folds* para encontrar la mejor combinación de parámetros. Por defecto, se toma $\gamma = 1/\text{atributos } (a)$ y $C = 1$.
- Random Forest: numTrees=100, seed=1. En algunos casos, se realiza una búsqueda de los parámetros maxDepth y numFeatures. Por defecto, se tiene maxDepth=unlimited, numFeatures=atributos (a),
- kNN: distance function=Euclidean distance, maxIterations=5.000, numClusters=3, seed=10.
- EM: maxIterations=1.000, minStdDev=1.0E-6, numClusters=3, seed=100.

En la Tabla 6.10 y 6.11 se recogen los errores de clasificación con el conjunto de datos del Modelo 1 y 2, respectivamente; en la Tabla 6.13 los errores correspondientes a algoritmos de *clustering* para esos dos modelos. En la Tabla 6.12 se muestran los errores obtenidos con algoritmos de clasificación para cada submodelo de la selección de atributos; y en la Tabla 6.14, los errores generados mediante técnicas de *clustering*. Dentro de cada una de estas tablas existe una fila para cada algoritmo utilizado en cada caso.

Tabla 6.10: Resultados de los algoritmos de clasificación del Modelo 1 (Falsos predictores). La primera columna indica el clasificador utilizado; las siguientes corresponden a las medidas de error.

Clasificador	Modelo 1				
	ErrClassif (%)	Prec	Recall	F-measure	AUC
Naive Bayes	2,57	0,975	0,974	0,974	0,994
J48	0,00	1,000	1,000	1,000	1,000
Multilayer Perceptron	0,00	1,000	1,000	1,000	1,000
SVM	3,06	0,970	0,969	0,970	0,993
Random Forest	7,89	0,921	0,921	0,920	0,978

Tabla 6.11: Resultados de los algoritmos de clasificación del Modelo 2 (Atributos predictores). La primera columna indica el clasificador utilizado; las siguientes corresponden a las medidas de error.

Clasificador	Modelo 2				
	ErrClassif (%)	Prec	Recall	F-measure	AUC
Naive Bayes	34,96	0,683	0,650	0,659	0,808
J48	27,99	0,693	0,720	0,698	0,799
Multilayer Perceptron	26,90	0,721	0,731	0,725	0,844
SVM	31,17	0,658	0,688	0,657	0,718
Random Forest	27,13	0,709	0,729	0,714	0,848

Tabla 6.12: Resultados de los algoritmos de clasificación del Modelo 3 (Atributos predictores tras selección) para cada submodelo. La primera columna indica el clasificador utilizado y las siguientes corresponden a las medidas de error de cada algoritmo de selección de atributos. Aparece en rojo el mejor resultado y en negrita los dos siguientes mejores para el submodelo seleccionado.

Modelo 3 - ReliefF					
Clasificador	ErrClassif (%)	Prec	Recall	F-measure	AUC
Naive Bayes	32,61	0,633	0,674	0,635	0,821
J48	32,26	0,640	0,677	0,631	0,782
Multilayer Perceptron	30,27	0,688	0,697	0,686	0,842
SVM	32,47	0,673	0,675	0,615	0,716
Random Forest	36,14	0,605	0,639	0,615	0,770

Modelo 3 - mRMR					
Clasificador	ErrClassif (%)	Prec	Recall	F-measure	AUC
Naive Bayes	32,85	0,637	0,671	0,624	0,826
J48	31,28	0,674	0,687	0,664	0,775
Multilayer Perceptron	29,57	0,699	0,704	0,697	0,846
SVM	31,57	0,680	0,684	0,662	0,734
Random Forest	32,47	0,658	0,675	0,657	0,812

Modelo 3 - GainRatioAttributeEval					
Clasificador	ErrClassif (%)	Prec	Recall	F-measure	AUC
Naive Bayes	27,54	0,700	0,725	0,703	0,806
J48	27,41	0,695	0,726	0,687	0,781
Multilayer Perceptron	27,50	0,694	0,725	0,694	0,815
SVM	27,79	0,674	0,722	0,665	0,745
Random Forest	27,01	0,704	0,730	0,700	0,812

Modelo 3 - InfoGainAttributeEval					
Clasificador	ErrClassif (%)	Prec	Recall	F-measure	AUC
Naive Bayes	32,93	0,700	0,671	0,678	0,817
J48	26,84	0,705	0,732	0,699	0,790
Multilayer Perceptron	25,19	0,735	0,748	0,738	0,866
SVM	26,11	0,722	0,739	0,710	0,765
Random Forest	30,65	0,676	0,694	0,683	0,815

Modelo 3 - CorrelationAttributeEval					
Clasificador	ErrClassif (%)	Prec	Recall	F-measure	AUC
Naive Bayes	29,09	0,696	0,709	0,700	0,832
J48	27,38	0,692	0,726	0,683	0,789
Multilayer Perceptron	25,16	0,736	0,748	0,738	0,867
SVM	26,59	0,712	0,734	0,702	0,762
Random Forest	27,79	0,693	0,722	0,694	0,836

Tabla 6.13: Resultados de los algoritmos de *clustering* del Modelo 1 (Falsos predictores) y Modelo 2 (Atributos predictores). La primera columna indica el *cluster* utilizado; las siguientes corresponden al error de clasificación de cada modelo de paciente.

	ErrClassif (%)	
Cluster	Modelo 1	Modelo 2
kNN	28, 16	44, 72
EM	47, 73	49, 58

Tabla 6.14: Resultados de los algoritmos de *clustering* del Modelo 3 (Atributos predictores tras selección) para cada submodelo. La primera columna indica el *cluster* utilizado; las siguientes corresponden al error de clasificación de cada algoritmo de selección de atributos: ReliefF, mRMR, *GainRatioAttributeEval*, *InfoGainAttributeEval* y *CorrelationAttributeEval*, respectivamente.

	Modelo 3 - ErrClassif (%)				
Cluster	ReliefF	mRMR	GainRatioAttr	InfoGainAttr	CorrelationAttr
kNN	39, 91	46, 95	44, 65	46, 94	47, 35
EM	60, 35	42, 63	37, 40	42, 67	37, 46

En la Tabla 6.10 se incluye el error de clasificación obtenido de los diferentes clasificadores con el modelo de paciente 1 (se recuerda que este conjunto no es representativo para un problema real, ya que incluye variables que son falsos predictores). Se puede observar que en la mayoría de los algoritmos utilizados el error es bajo (cercano a cero), debido a que el modelo aprende con ciertas variables que se han denominado de resultado en el estado del arte, las cuales aparecen tras el parto y tienen ya información inherente sobre la clase. Es importante destacar que los errores obtenidos con SVM y Random Forest no son tan favorables como se esperaba debido a que por limitaciones de tiempo no se ha llegado a obtener una combinación de parámetros óptima.

Sin embargo, no ocurre lo mismo cuando se eliminan las variables de resultado del conjunto, y se entrena el modelo sin dicha información. En la Tabla 6.11 se pueden observar que los resultados mejoran el error inicial del modelo de referencia. En este caso, se obtienen los mejores resultados (26, 90 % de error) con Multilayer Perceptron.

El objetivo de este proyecto es determinar un subconjunto de variables relevantes que permitan mejorar los resultados de los modelos anteriores. Por ello, en la Tabla 6.12 se recogen las medidas de error resultantes de aplicar los cinco clasificadores a cada submodelo obtenido de la selección de atributos. En ella, se muestran los resultados del Modelo 3, el cual contiene un subconjunto de variables predictoras, presentando una notable mejora frente a los modelos anteriores. Los errores en este caso son similares en casi todos los algoritmos y submodelos de

paciente, llegando a alcanzar el mejor error, un 25,16 % (marcado en rojo), con el clasificador Multilayer Perceptron y el submodelo de variables obtenido con *CorrelationAttributeEval*. Junto a este clasificador también se marcan en negrita los dos algoritmos (J48 y SVM) que mejores resultados proporcionan dentro de este submodelo en términos de error de clasificación. Los peores resultados en esta ocasión coinciden con el clasificador Naive Bayes, lo cual conduce a pensar que los atributos que componen cada submodelo no son independientes y, es por ello, que este clasificador presente peores resultados.

También se han analizado los errores obtenidos mediante técnicas de *clustering*. En las Tablas 6.13 y 6.14 se observa el mismo fenómeno que ocurre en las técnicas de clasificación, que el Modelo 1 proporciona mejores resultados debido a que contienen variables correlacionadas con la clase y, por tanto, no se pueden utilizar para un problema real. Los resultados de *clustering* de los Modelos 2 y 3 son peores que los obtenidos en clasificación.

6.4.3. Evaluación comparativa

A partir de estos resultados se ha observado que es posible seleccionar un modelo de paciente que permita obtener mejores resultados del modelo de referencia, pasando de un 41,89 % a un 25,16 % de error.

Para concluir, la solución que se empleará en la herramienta en un funcionamiento real corresponde al Modelo 3 con los seis atributos *motivo_induccion*, *rpm_ingreso*, *bishop_ingreso*, *motivo_cesarea_previa*, *cesareas_previas* y *partos_vaginales_previos* y el clasificador Multilayer Perceptron. Se observa también que cuatro de esas variables el experto las consideraba con peso máximo (véase Tabla 4.1), es decir, las más importantes para determinar si hay inducción; y las dos restantes con el segundo peso más alto.

Además, se puede apreciar cierta relación entre la complejidad del modelo de paciente y los algoritmos que mejor funcionan. Esto es, para los submodelos más simples obtenidos del Modelo 2 tras la selección de atributos, en media, funcionan mejor los clasificadores (en concreto, la red neuronal) frente a modelos algo más complejos como son el propio Modelo 2, como se muestra en la Figura 6.1. Sin embargo, se descarta el Modelo 1 del sistema, ya que no es representativo del problema real estudiado, lo cual se ve apoyado por los errores medidos.

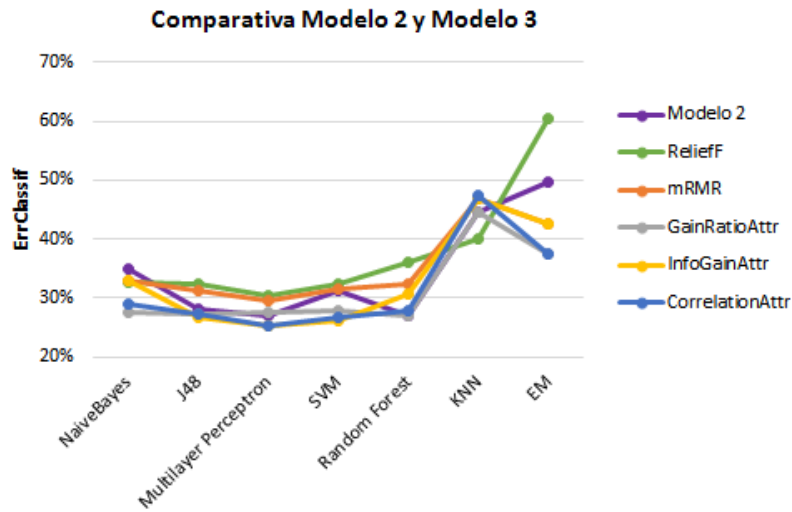


Figura 6.1: Comparativa del error de clasificación entre los Modelos 2 y cada submodelo del Modelo 3 para los distintos algoritmos utilizados.

En las matrices de confusión de los clasificadores probados (Anexo D) se observa que trabajar con clases desbalanceadas puede implicar un sesgo en los clasificadores, priorizando la clase mayoritaria y realizando una mala clasificación de las instancias de las clases minoritarias. Para solventar este problema se proponen diferentes estrategias en la literatura [37] agrupadas en las siguientes categorías:

- Nivel de datos: re-muestreo de la base de datos para balancear las clases. Consiste en alcanzar un balance entre las clases mediante la eliminación de instancias de la clase mayoritaria (sub-muestreo) o la inclusión de objetos en la clase minoritaria (sobre-muestreo). El sub-muestreo puede excluir algunos objetos representativos o valiosos para entrenar el clasificador. En cambio, el sobre-muestreo incluye objetos artificiales que pueden llevar a *overfitting*.
- Modificación de algoritmos: los clasificadores existentes son modificados para fortalecer su predicción con respecto a la clase minoritaria. Depende mucho de la naturaleza del clasificador y la mayoría son modificados para resolver un problema específico.
- Aprendizaje sensitivo al coste: considera los costes asociados a las instancias mal clasificadas. Si el conjunto de datos no está balanceado, se asigna un mayor peso a los errores sobre la clase minoritaria y se desarrolla una hipótesis que minimice el coste total sobre el conjunto de entrenamiento.

Como línea futura de investigación se propone estudiar las estrategias anteriores para mitigar las dificultades que aparecen en la clasificación supervisada al trabajar con datos con clases desbalanceadas. Además de probar varios clasificadores con sus respectivos parámetros óptimos

y reasignar los pesos asociados a cada una de las variables del conjunto de datos en base a los resultados obtenidos.

6.5. Explicación en clasificación

El objetivo fundamental de esta subsección es proporcionar una explicación de la clasificación obtenida a partir del sistema de aprendizaje automático. Para ello, se ha implementado el Algoritmo 1, explicado en el estado del arte, el cual se caracteriza por dotar de transparencia a los modelos ofreciendo así ventajas tanto a expertos en modelos como a cualquier usuario que utilice el sistema.

Los algoritmos entrenados son la red neuronal (Multilayer Perceptron), la máquina de vectores soporte (SVM) y el árbol de decisión (J48), ya que son los tres, en ese orden, con mejor error de clasificación en el submodelo de paciente obtenido con *CorrelationAttributeEval* del Modelo 3. El último de estos algoritmos entrenados muestra las explicaciones ofrecidas y, gracias a su transparencia, es posible contrastar con el propio árbol. Por el contrario, los otros dos muestran el verdadero potencial del método implementado, ya que proporciona explicaciones de las predicciones de modelos considerados de *caja negra*.

Este experimento se lleva a cabo sobre el submodelo de paciente antes mencionado, del que se seleccionan únicamente cuatro de sus seis atributos, para simplificar el análisis y la visualización de los resultados. De manera que el conjunto de datos empleado se compone de 10.487 instancias con los atributos de tipo numérico y nominal siguientes: *motivo_induccion*, *rpm_ingreso*, *bishop_ingreso* y *motivo_cesarea_previa*. La variable objetivo (*resultado*) es la decisión elegida antes del inicio del parto y, como se ha explicado en el documento, toma los tres posibles valores *No inducción*, *Inducción* y *Cesárea*, los cuales serán referenciados en las gráficas como clase 1, 2 y 3, respectivamente. En la Figura 6.2 se representa la proyección del conjunto de datos sobre los subespacios definidos por cada par de atributos considerado en el estudio. En ella se puede apreciar la complejidad del mismo, pues no resulta tan fácil separar las tres clases.

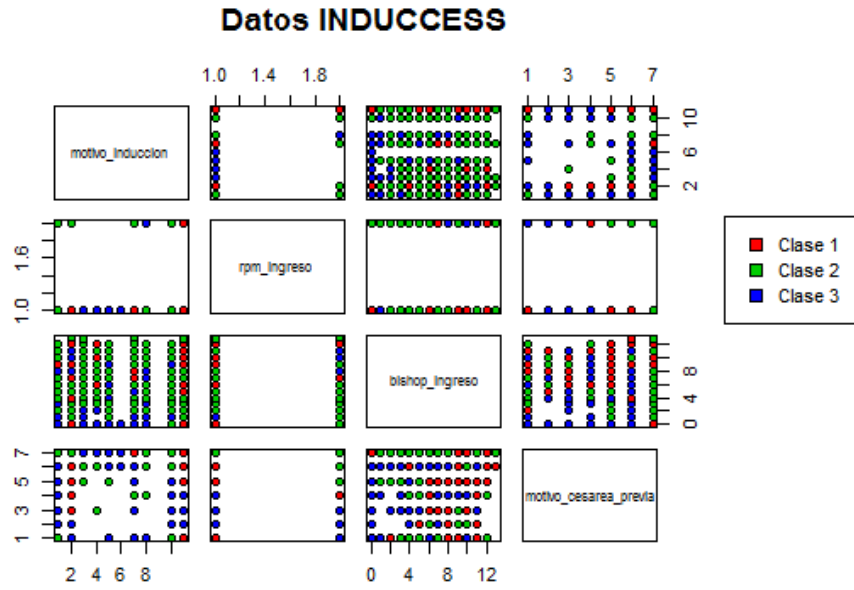
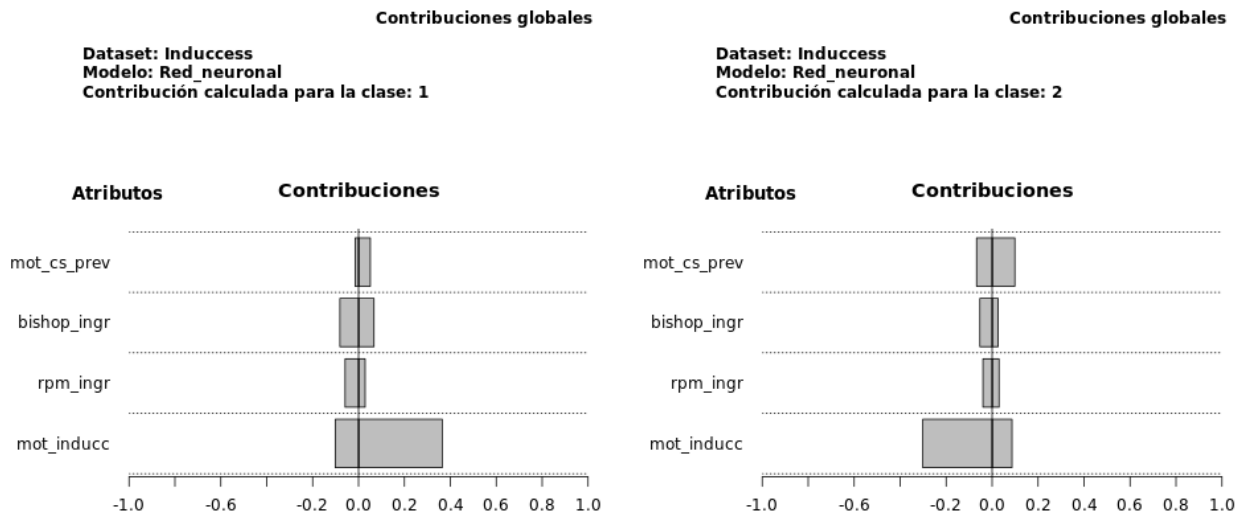


Figura 6.2: Representación bidimensional del conjunto de datos INDUCESS. En cada cuadrado se sitúan las instancias del conjunto en función de los valores que toman los dos atributos representados. Se muestran de diferente color las clases a las que pertenecen las instancias.

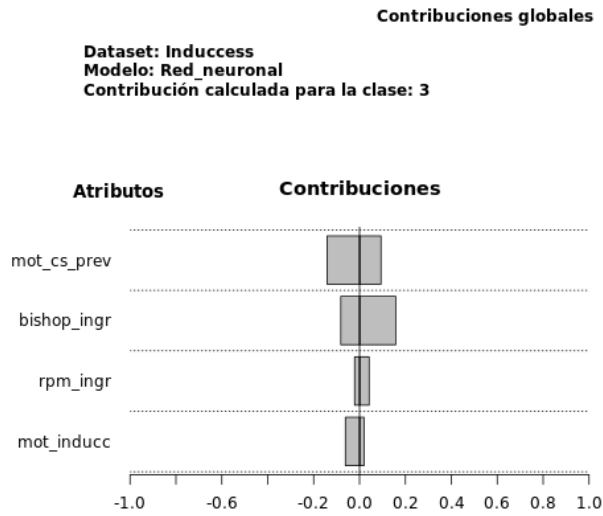
En la implementación del Algoritmo 1, al igual que en los experimentos anteriores, se realiza *cross-validation* de 10 *folds* para la obtención de las predicciones de todas las instancias. Además, las explicaciones de las predicciones individuales se promedian para obtener las contribuciones de un valor (o rango de valores) de un atributo determinado y, a su vez, promediando éstas se generan las contribuciones globales del atributo hacia la predicción de la clase. A la hora de promediar, se consideran por separado las contribuciones positivas y las negativas, ya que, de lo contrario, la contribución de un valor o de un atributo puede resultar casi nula cuando en realidad es muy influyente en ambos sentidos.

A continuación, se presentan varias gráficas sobre explicaciones generadas en diversas situaciones. En primer lugar las obtenidas con la red neuronal, puesto que fue el algoritmo ganador en el estudio anterior, seguidas de las obtenidas con los dos siguientes mejores algoritmos, esto es, SVM y árbol. En cada una de esas gráficas se especifica si las explicaciones corresponden a las contribuciones de los valores de un atributo o a las contribuciones globales de todos los atributos.



(a) Contribuciones a la predicción de la clase *No inducción*.

(b) Contribuciones a la predicción de la clase *Inducción*.



(c) Contribuciones a la predicción de la clase *Cesárea*.

Figura 6.3: Contribuciones globales de los atributos *motivo_cesarea_previa*, *bishop_ingreso*, *rpm_ingreso* y *motivo_induccion* a la predicción de cada clase para una red neuronal.

En la Figura 6.3 se incluyen las contribuciones globales de los cuatro atributos que se consideran en el conjunto de datos en la predicción de una red neuronal. Se puede observar en los resultados que, dependiendo de la clase a predecir, ciertos atributos tienen más influencia que otros. Concretamente, para predecir la clase *No inducción* y clase *Inducción* el atributo más influyente en la decisión del modelo es *motivo_induccion*, aunque en el primer caso de manera positiva y en el segundo negativamente. Dicho atributo corresponde a un indicador en el que el experto se apoya para determinar si se realiza una inducción al parto. En la información extraída de la experiencia de los profesionales, ya se había visto que esta variable es de las más importantes en la decisión y, por ello, en la Tabla 4.1 aparece con peso 5-6. Sin embargo, este atributo presenta contribuciones fuertemente positivas para determinar los casos de *No inducción* (clase 1). Si se observa el histograma de valores que toma el atributo (Figura 6.4) existe un alto número de instancias que pertenecen a la clase *No inducción* (clase 1, representada en rojo) con valor *NA*, lo que indica que el hecho de no existir ningún motivo de inducción en la Historia Clínica del paciente conduce a la decisión de no inducir. Este razonamiento es apoyado por lo que refleja la imagen superior derecha que recoge las contribuciones para predecir la clase 2 (*Inducción*), donde el atributo *motivo_induccion* influye negativamente y, por tanto, descarta la pertenencia a esta clase. Para la clase 2, aunque no con tanta claridad como en la clase 1, destaca como atributo influyente *motivo_cesarea_previa*. Finalmente, para predecir la clase 3 (*Cesárea*) con esta red neuronal, se puede observar que todos los atributos tienen cierta influencia aunque las contribuciones son bajas, destacando con valores positivos *bishop_ingreso*.

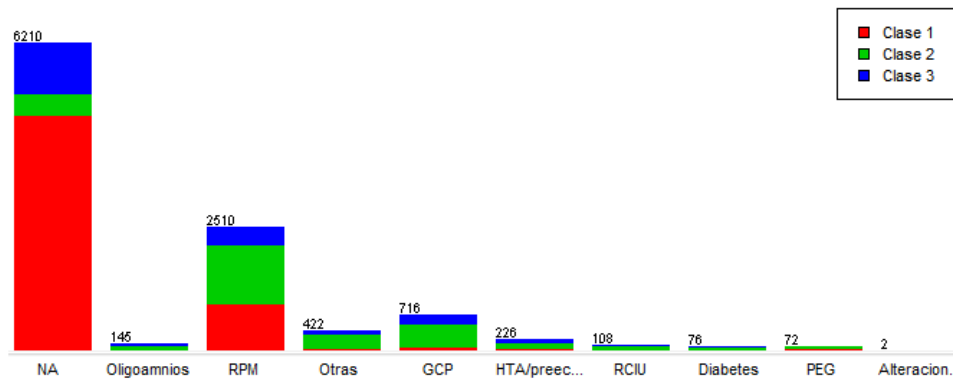
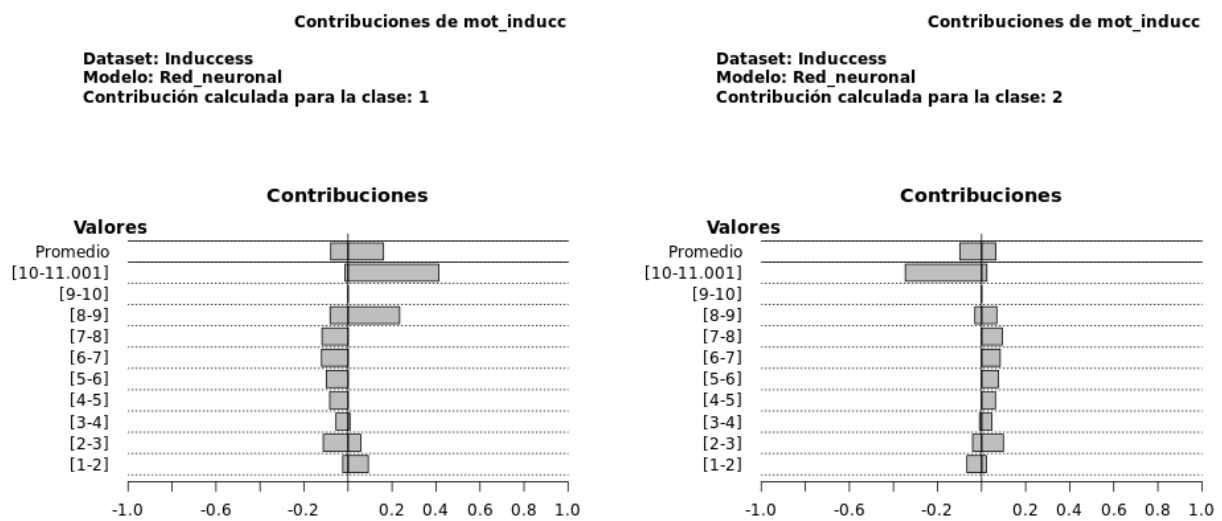
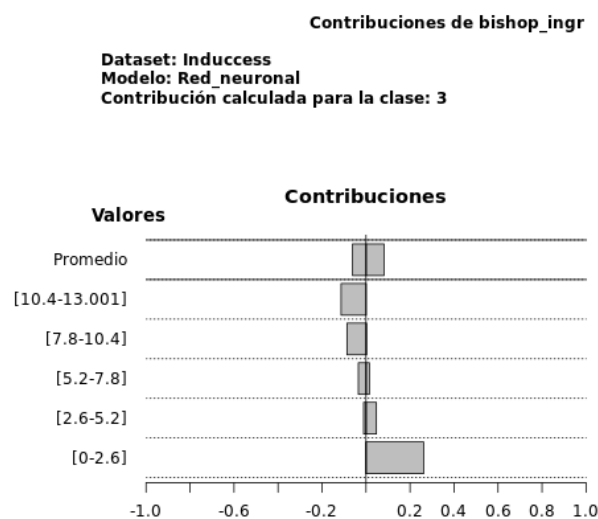


Figura 6.4: Histograma de *motivo_induccion* en el conjunto de datos INDUCESS. Aparece el valor que toma el atributo en el eje *x*, en color rojo la clase 1 (*No inducción*), en verde la clase 2 (*Inducción*) y en azul la clase 3 (*Cesárea*).



(a) Contribuciones de *motivo_induccion* a la predicción de la clase *No inducción*. El eje *y* muestra los valores 1:*GCP*, 2:*RPM*, 3:*RCIU*, 4:*PEG*, 5:*Oligoamnios*, 6:*Alteración bienestar fetal*, 7:*HTA / preeclampsia / eclampsia / EHE*, 8:*Diabetes*, 9:*Patología médica materna*, 10:*Otras*, 11:*NA*.

(b) Contribuciones de *motivo_induccion* a la predicción de la clase *Inducción*. El eje *y* muestra los valores 1:*GCP*, 2:*RPM*, 3:*RCIU*, 4:*PEG*, 5:*Oligoamnios*, 6:*Alteración bienestar fetal*, 7:*HTA / preeclampsia / eclampsia / EHE*, 8:*Diabetes*, 9:*Patología médica materna*, 10:*Otras*, 11:*NA*.



(c) Contribuciones de *bishop_ingreso* a la predicción de la clase *Cesárea*. El eje *y* muestra un rango numérico.

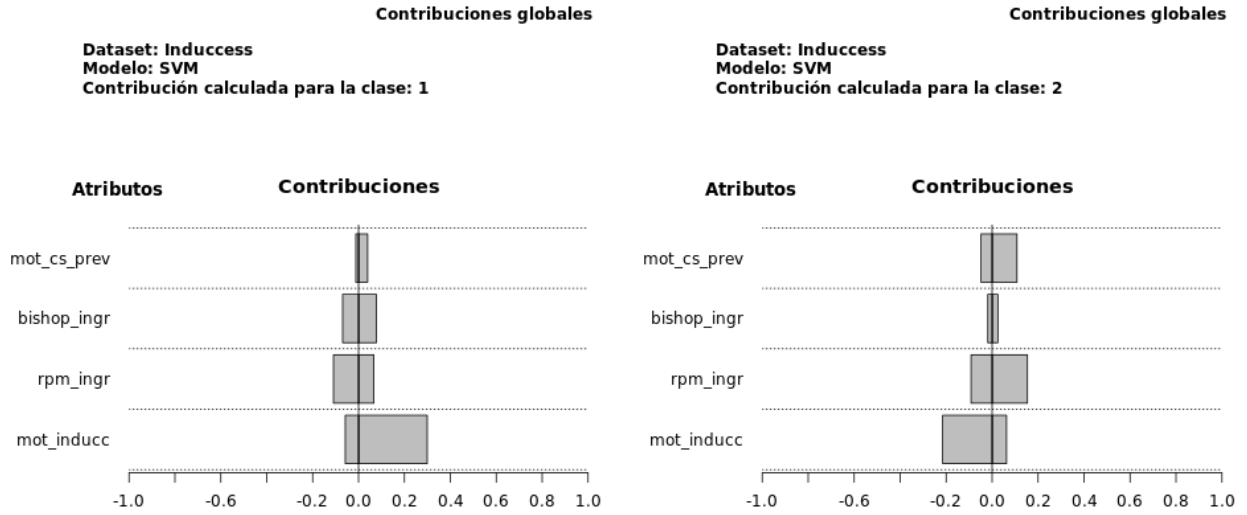
Figura 6.5: Contribuciones de diferentes valores de un atributo a la predicción de las clases para una red neuronal. La primera barra (*Promedio*) representa el promedio de las contribuciones de todos los rangos de valores.

En la Figura 6.5 se recogen las contribuciones de diferentes rangos de valores de atributos representativos en la predicción de cada clase, así como el promedio de las positivas y negativas para todos ellos. En cada caso, se ha establecido una distribución en rangos distintos dependiendo de los valores que tome, ya que el programa implementado permite especificar cualquier número de rangos en los que dividir el intervalo de posibles valores de cada atributo. Las dos primeras imágenes muestran las contribuciones del atributo *motivo_induccion* para las clases *No inducción* e *Inducción*, respectivamente. Ambas reflejan que para instancias con valores muy elevados de este atributo son asignadas a la clase *No inducción* (10:Otras y 11:NA e, incluso, 8:Diabetes), mientras que en valores intermedios es más probable que la red neuronal asigne la clase *Inducción*. Sin embargo, en la imagen inferior se observa que para determinar las instancias de la clase *Cesárea* la zona de valores pequeños del atributo *bishop_ingreso* son los más influyentes, coincidiendo justamente con una de las reglas de decisión proporcionadas por el experto en el campo (*‘Si bishop_ingreso ≤ 6 entonces Cesárea’*).

Por otro lado, las contribuciones globales de los atributos hacia la predicción de las tres clases mediante máquinas de vectores soporte se incluyen más abajo en la Figura 6.6. En este caso, aunque con contribuciones algo más bajas que en la red neuronal, se observa que para las clases *No inducción* y *Cesárea* el atributo influyente coincide, es decir, de nuevo destacan *motivo_induccion* en la clase *No inducción* y *bishop_ingreso* en la clase *Cesárea*. Por el contrario, para predecir la clase *Inducción* no se obtiene lo que sucede con la red neuronal y, en este caso, es *rpm_ingreso* el atributo más influyente, a pesar de que *motivo_cesarea_previa* también presenta valores positivos.

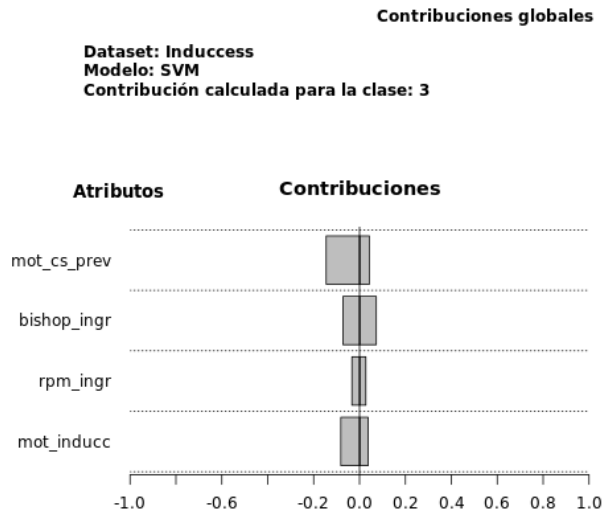
Con el objetivo de determinar cuál de los dos atributos es más relevante en la predicción de la clase 2 (*Inducción*), en la Figura 6.7 se muestran las explicaciones por rangos de valores tanto para la SVM, en las imágenes de la izquierda, como para la red neuronal, en las de la derecha. En el caso de *rpm_ingreso*, las imágenes superiores reflejan que la mayoría (todas en la SVM) de las instancias de la clase *Inducción* están confinadas en valores altos del atributo, que coincide con la opción *Sí* del mismo. Sin embargo, las bajas contribuciones en el caso de la red neuronal pueden indicar que este atributo no sea realmente influyente para determinar esta clase o, por el contrario, sea el modelo el que no haya capturado bien el dominio del problema, ya que como se ha visto, anteriormente se obtuvo *motivo_cesarea_previa*. No obstante, la existencia de contribuciones tanto positivas como negativas para los dos rangos de valores, pone de manifiesto la extrema complejidad del problema debido a la influencia de muchos otros factores en la decisión. En las imágenes inferiores se muestra que *motivo_cesarea_previa* también es un atributo influyente y en ambos algoritmos es más probable que se asigne la clase 2 (*Inducción*) a una instancia para valores altos de este atributo (5:Podálica, nalgas, 6:Otras, 7:NA). La mayoría de las instancias de la clase 2 toman un valor de *NA* en este atributo, lo que conlleva a pensar que, aunque del conocimiento extraído del experto se ha visto que esta variable se utilice para determinar una cesárea, el hecho de no aparecer información de la misma

en una instancia (valor *NA*) descarta predecir la clase 3 (*Cesárea*), pudiendo ser cualquiera de las restantes y, en este caso, la 2.



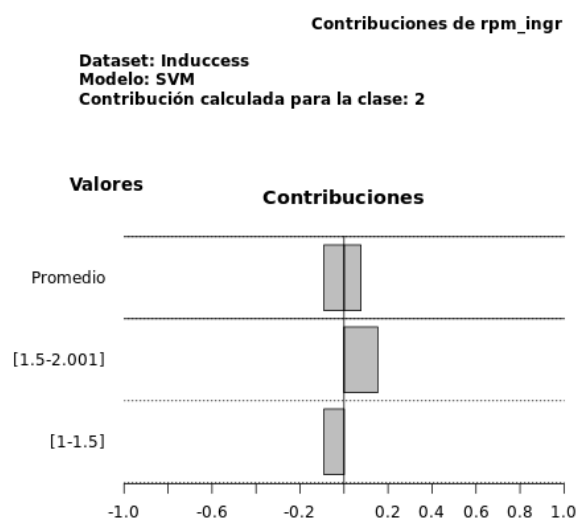
(a) Contribuciones a la predicción de la clase *No inducción*.

(b) Contribuciones a la predicción de la clase *Inducción*.

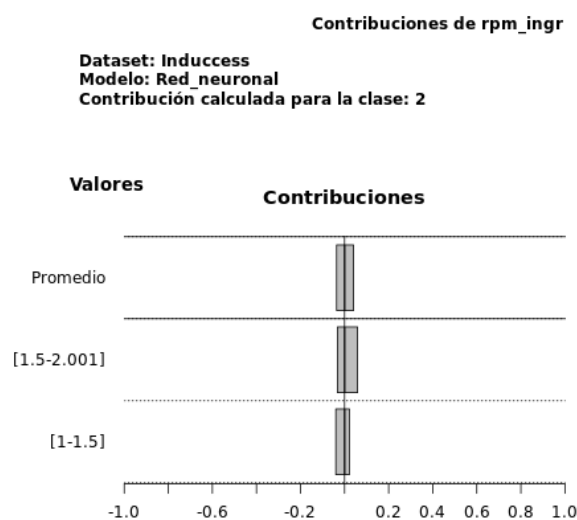


(c) Contribuciones a la predicción de la clase *Cesárea*.

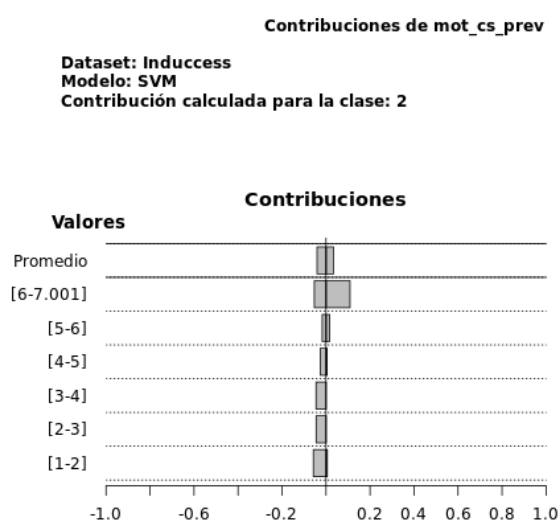
Figura 6.6: Contribuciones globales de los atributos *motivo_cesarea_previa*, *bishop_ingreso*, *rpm_ingreso* y *motivo_induccion* a la predicción de cada clase para una SVM.



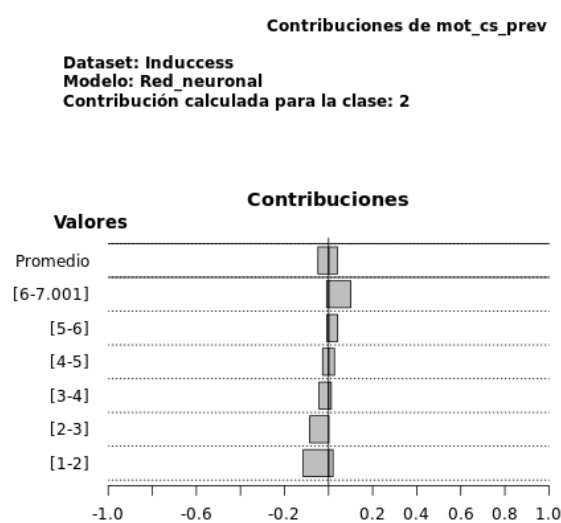
(a) Contribuciones de *rpm_ingreso* a la predicción de la clase *Inducción* para una SVM. El eje *y* muestra los valores 1:No, 2:Sí.



(b) Contribuciones de *rpm_ingreso* a la predicción de la clase *Inducción* para una red neuronal. El eje *y* muestra los valores 1:No, 2:Sí.

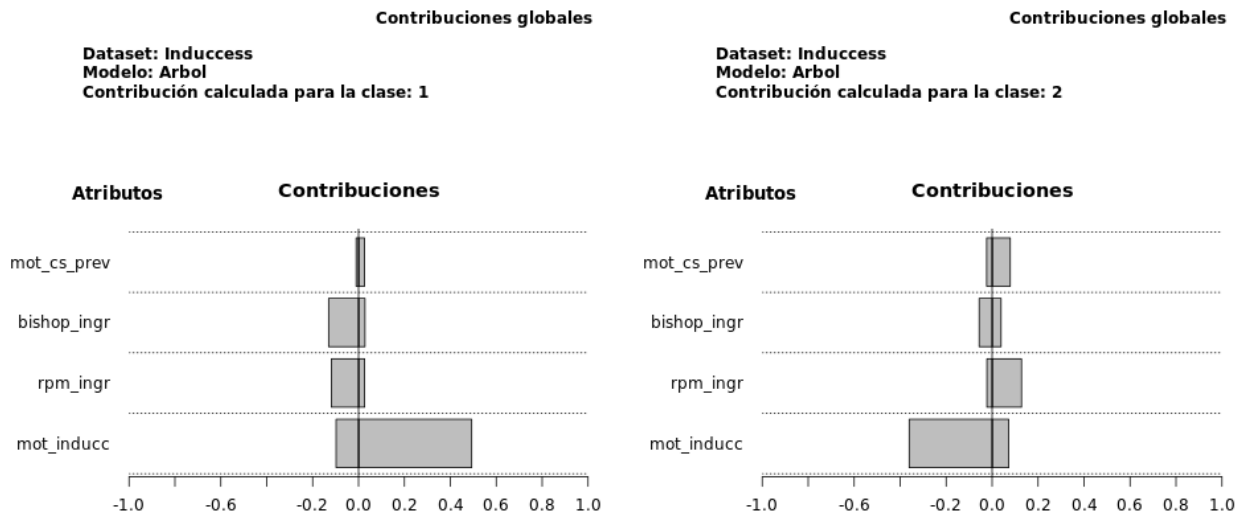


(c) Contribuciones de *motivo_cesarea_previa* a la predicción de la clase *Inducción* para una SVM. El eje *y* muestra los valores 1:RPBF (*riesgo pérdida bienestar fetal*), 2:FI (*fracaso inducción*), 3:NPP (*no progresión parto*), 4:DPC (*desproporción pélvico-cefálica*), 5:Podálica, nalgas, 6:Otras, 7:NA.



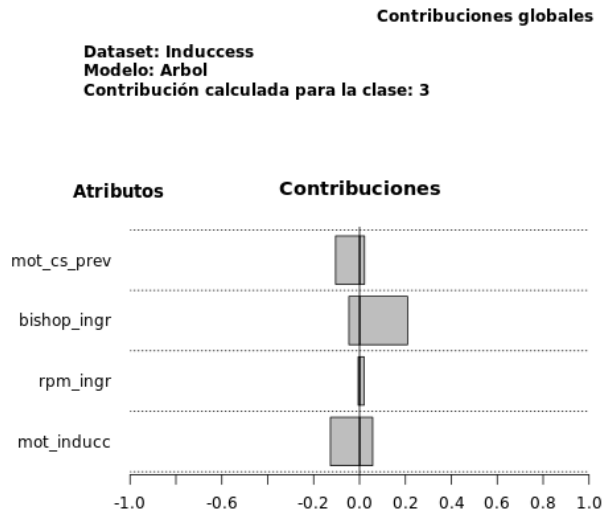
(d) Contribuciones de *motivo_cesarea_previa* a la predicción de la clase *Inducción* para una red neuronal. El eje *y* muestra los valores 1:RPBF (*riesgo pérdida bienestar fetal*), 2:FI (*fracaso inducción*), 3:NPP (*no progresión parto*), 4:DPC (*desproporción pélvico-cefálica*), 5:Podálica, nalgas, 6:Otras, 7:NA.

Figura 6.7: Contribuciones de diferentes valores de un atributo a la predicción de la clase *Inducción*. La primera barra (*Promedio*) representa el promedio de las contribuciones de todos los rangos de valores.



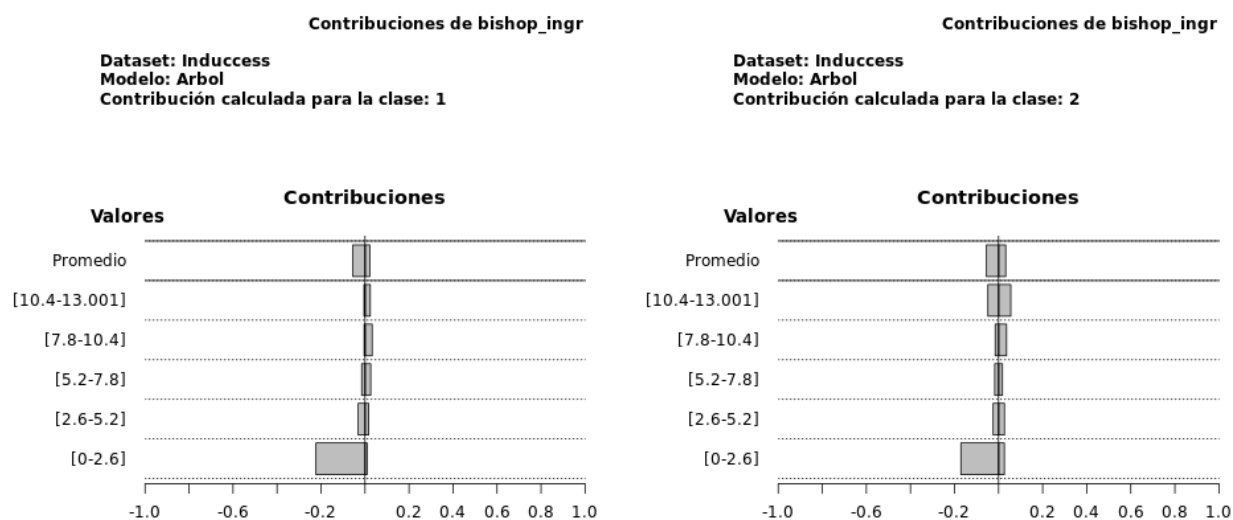
(a) Contribuciones a la predicción de la clase *No inducción*.

(b) Contribuciones a la predicción de la clase *Inducción*.



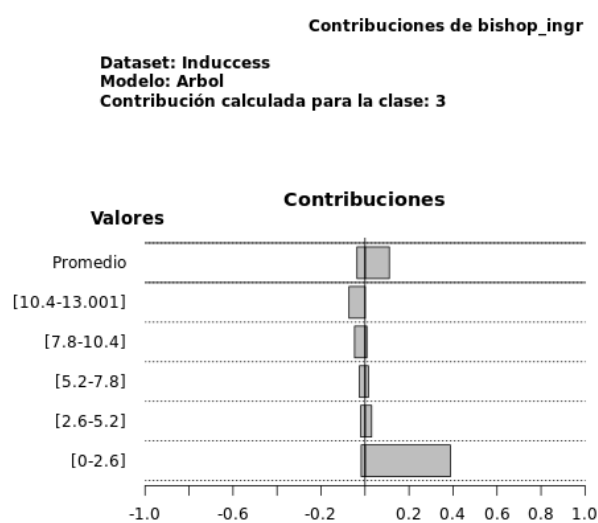
(c) Contribuciones a la predicción de la clase *Cesárea*.

Figura 6.8: Contribuciones globales de los atributos *motivo_cesarea_previa*, *bishop_ingreso*, *rpm_ingreso* y *motivo_induccion* a la predicción de cada clase para un árbol.



(a) Contribuciones de *bishop_ingreso* a la predicción de la clase *No inducción*. El eje *y* muestra un rango numérico.

(b) Contribuciones de *bishop_ingreso* a la predicción de la clase *Inducción*. El eje *y* muestra un rango numérico.



(c) Contribuciones de *bishop_ingreso* a la predicción de la clase *Cesárea*. El eje *y* muestra un rango numérico.

Figura 6.9: Contribuciones de diferentes valores de un atributo a la predicción de las clases para un árbol. La primera barra (*Promedio*) representa el promedio de las contribuciones de todos los rangos de valores.

En la Figura 6.8 se puede ver que las contribuciones globales obtenidas con un árbol de decisión son similares a las obtenidas con una SVM, aunque en este caso los valores positivos están más pronunciados, concretamente en la predicción de la clase *Cesárea*, como ocurre en la red neuronal. Pero, en todos los casos, coincide que el atributo más influyente es *bishop_ingreso*.

Si se comparan las contribuciones por rangos de valores del atributo citado hacia la predicción de cada clase, se observa claramente (Figura 6.9) que para valores pequeños del atributo las contribuciones son positivas para la clase *Cesárea* y negativas en el resto, descartando la pertenencia a las clases *No inducción* e *Inducción*. Además, la gráfica superior derecha refleja, de manera no tan clara, que para valores más altos del atributo se podría asignar la clase *Inducción*, aunque la clase *No inducción* tampoco se descarta. Como ya se ha comentado anteriormente, estas explicaciones coinciden con la práctica clínica, pues entre las reglas de decisión del sistema experto se encuentran las dos siguientes: ‘*Si bishop_ingreso* ≤ 6 entonces *Cesárea*’ y ‘*Si bishop_ingreso* > 6 entonces *Inducción*’.

Es importante remarcar que las contribuciones anteriores reflejan lo que ha aprendido el modelo entrenado. No hacen referencia directa a la verdadera distribución de las instancias en el espacio de atributos. Lo que sucede en algunos de las situaciones descritas es que los modelos entrenados han capturado bien el dominio real del problema y las contribuciones, además de explicar el funcionamiento del modelo, reflejan con bastante fidelidad dicho ámbito.

Tras observar los resultados obtenidos y el análisis realizado, se considera que se ha cumplido con una de las motivaciones principales del proyecto, de valorar la relevancia de las variables del modelo de referencia y encontrar otros modelos que determinen aquellas realmente relevantes y mejoren los resultados a la hora de predecir el éxito de una inducción del parto. A la vez que se incorporan técnicas de visualización e interpretación de resultados más sencillas e intuitivas para cualquier usuario experto o no en el campo del aprendizaje automático.

7

Conclusiones y trabajo futuro

7.1. Conclusiones

El trabajo presentado se puede dividir en tres bloques: una primera parte, en la que se ha logrado diseñar un sistema complejo que permite explotar la información sobre embarazos recogidos en el Hospital Universitario de Fuenlabrada y, con ello, extraer valor de los datos; una segunda parte, en la que se desarrolla un sistema que permite inferir el éxito de la inducción al parto. Se construye además una herramienta de ayuda a la toma de decisiones basada en reglas de decisión provenientes de un panel de expertos en Obstetricia y se evalúan diferentes técnicas y algoritmos de aprendizaje automático y Big Data; y una última parte, donde se lleva a cabo el análisis de una técnica que permite dotar de transparencia los modelos considerados.

En primer lugar, el diseño del sistema comienza describiendo el modelado de los datos, así como las fases necesarias para abordarlo. Se han diseñado con éxito herramientas para la recopilación masiva de información clínica así como un sistema de validación y visualización de los datos. Habida cuenta de los resultados obtenidos tras la validación, aún queda trabajo por realizar a fin de minimizar la posible discrepancia entre el modelo de paciente y la información real y consolidar una base de datos robusta.

En segundo lugar, la decisión sobre iniciar un mecanismo para finalizar exitosamente un parto no es un tema sencillo de abordar y de resolver. Por ello, hasta ahora, no existen herramientas que ayuden al experto en Obstetricia en la decisión. Se ha logrado desarrollar un sistema de ayuda a la toma de decisiones basado en reglas, después de pasar por varias etapas intermedias de revisión y reformulación de reglas, las cuales logran recoger y manejar gran parte del conocimiento y experiencia del experto. El sistema diseñado integra todos los datos recopilados de

los pacientes para obtener la decisión final, y, por tanto, se considera una aproximación real al diagnóstico que elabora el experto.

Por otra parte, se han revisado los principales modelos de aprendizaje automático, valorando cualitativamente su adecuación para ser aplicados a problemas del ámbito médico. A fin de obtener modelos de mayor riqueza de información y evitar variables que lastren los resultados, se han estudiado técnicas de selección de atributos, de las cuales se obtienen otros modelos de datos de paciente sobre los que realizar las pruebas. La comparativa de todos estos modelos ha resultado ser muy positiva y permite concluir que existe cierta información que introduce ruido en los modelos y, sin ella, los resultados mejoran notablemente. Cabe destacar que para obtener estas conclusiones se ha realizado un exhaustivo análisis de los modelos y clasificadores candidatos y, por ello, se tiene un alto nivel de confianza en la validez de las conclusiones presentadas.

En último lugar, tras seleccionar el modelo de paciente y clasificadores adecuados, se da continuidad y se amplía una vía de investigación para explicar la clasificación obtenida. Se ha estudiado la técnica que permite conocer cómo contribuye el valor de cada atributo y el propio atributo en sí a la predicción. Para ello, se ha implementado el algoritmo y un correspondiente sistema apropiado para visualizar los resultados.

Tratando los diferentes aspectos del proyecto más en detalle, se pueden extraer algunas conclusiones más.

En primer lugar, se ha observado que la extracción y el modelado de información es una etapa clave a la hora de desarrollar un sistema de inferencia del éxito de inducción del parto. De hecho, se han encontrado ciertos atributos cuya información no resultaba útil o eran realmente redundantes en el modelo sesgando los resultados.

En segundo lugar, se han estudiado y propuesto otros modelos de paciente y algoritmos de aprendizaje automático con mayor riqueza que el modelo de referencia, con el fin de medir la efectividad del mismo. Todo este trabajo de análisis de atributos concluirá en la creación de un modelo de paciente más apropiado, y por ende, que mejore los resultados. Finalmente, a pesar del desbalanceo de clases, que limita cualquier método de clasificación, el error del modelo de referencia se ha visto reducido a un 25,16 % considerando únicamente seis atributos relevantes con una red neuronal. A consecuencia de esto, se puede concluir que resulta de vital importancia realizar un estudio de atributos antes de pasar a generar los modelos de paciente.

Adicionalmente, los resultados obtenidos tras la implementación y aplicación del algoritmo de explicación de las predicciones han sido muy satisfactorios, coincidiendo con la práctica clínica real. Las contribuciones calculadas explican cómo emite un modelo las predicciones a partir de los valores de los atributos. En el conjunto de datos INDUCESS se ve reflejado que, en general, para determinar *No inducción* basta con que no exista un motivo por el cual el experto podría recomendar inducción; para *Inducción* puede ser o, bien que no exista un motivo de cesárea

previa, es decir, no tenga ninguna cesárea previa, o bien que presente ruptura prematura de membranas (*rpm*) al ingreso; y para determinar una *Cesárea* suele ocurrir que el índice de Bishop al ingreso sea menor que 6. Estas explicaciones permiten hacer más transparentes los modelos considerados y quizás ayudar a complementar los conocimientos o descubrir relaciones entre los datos que han pasado desapercibidas hasta el momento. Todo ello supone una enorme ventaja para que sea utilizado por cualquier usuario.

El sistema completo implementado ha resultado ser de gran interés y de utilidad en la ayuda en la toma de decisiones para el experto, pudiendo ser aplicado a cualquier otra rama de la medicina. Mediante esta herramienta, se ha podido extraer y visualizar la riqueza que contiene la información clínica así como diferentes aspectos a tener en cuenta para modelar los pacientes. Pero sus posibilidades no se limitan a dichos aspectos, sino que permitirá probar y evaluar otros puntos del sistema como reglas de decisión, otras prioridades o diversos clasificadores. Finalmente, con este proyecto se abre una nueva línea de trabajo donde se va a llevar a cabo el desarrollo de la herramienta para ser utilizada en la práctica clínica en tiempo real por el personal médico experto.

7.2. Trabajo futuro

Existen múltiples líneas de continuación del trabajo realizado hasta este punto. En particular, se puede destacar una serie de cuestiones que han ido surgiendo en el desarrollo de este trabajo, que, bien por quedar fuera de los objetivos iniciales o bien por falta de tiempo, no han sido abordadas.

En primer lugar, se ha resaltado en el proyecto la importancia de la calidad de la información que se utilice en los modelos. Aunque el sistema de recuperación de información se ha realizado de manera bastante rigurosa, los resultados de validación han sido satisfactorios y se considera adecuado, se debería realizar alguna iteración más y revisar muestras más grandes de pacientes en el proceso, para así conseguir refinar el sistema lo máximo posible.

En segundo lugar, se considera un punto crítico la manera de abstraer conocimiento del experto y representarlo con un lenguaje entendible por el sistema. Por ello, se propone un cambio en el sistema experto, que consiste en desarrollarlo partiendo de las reglas provenientes de expertos con diferentes prioridades y, que a su vez, sea capaz de aprender con cada utilización y, con ello, realimentar continuamente su base de conocimiento. Otra opción sería aplicar un algoritmo genético para generar reglas automáticamente y así hacer más robusto el sistema experto.

Otra línea posible, sería mitigar las dificultades que aparecen en la clasificación cuando se trabaja con clases desbalanceadas. De modo que, una muestra balanceada permitiría refinar la clasificación y eliminar el sesgo que afectaba negativamente a la capacidad de extraer patrones

de la clase minoritaria.

Por otra parte, el principal objetivo de este proyecto era desarrollar un modelo predictivo del éxito de un procedimiento de inducción del trabajo de parto y la herramienta desarrollada es adecuada para ello. Se propone extenderla para que sea realmente operativa y ayude al médico experto en la toma de decisiones en su práctica diaria. Por tanto, será necesario ampliar la aplicación, y entrenar y probar con otras técnicas y algoritmos de aprendizaje automático. Para conseguir este nuevo objetivo no sólo sería conveniente disponer de conexión directa con la plataforma Selene sino que se debería automatizar todo el proceso descrito en el proyecto. De este modo y, gracias al éxito dentro del ámbito médico, el sistema podría ser utilizado en otras áreas como ayuda en la validación de informes redactados por profesionales.

Bibliografía

- [1] C. J. Verhoeven, A. Oudenaarden, M. A. Hermus, M. M. Porath, S. G. Oei, and B. W. J. Mol, “Validation of models that predict Cesarean section after induction of labor,” *Ultrasound in Obstetrics & Gynecology*, vol. 34, no. 3, pp. 316–321, 2009.
- [2] N. Baños, F. Migliorelli, E. Posadas, J. Ferreri, and M. Palacio, “Definition of Failed Induction of Labor and Its Predictive Factors: Two Unsolved Issues of an Everyday Clinical Situation,” *Fetal diagnosis and therapy*, 2015.
- [3] F. Almachi, S. Lisette *et al.*, “Eficacia y seguridad de la inducción del trabajo de parto con misoprostol en pacientes con embarazo a término con indicación de inducción y cervix desfavorable,” *Tesis Doctoral, Escuela de Obstetricia - Universidad de Guayaquil*, 2013.
- [4] W. A. Grobman, “Predictors of induction success,” in *Seminars in perinatology*, vol. 36, no. 5. Elsevier, 2012, pp. 344–347.
- [5] J. Di Deco Sampedro *et al.*, “Estudio y aplicación de técnicas de aprendizaje automático orientadas al ámbito médico: estimación y explicación de predicciones individuales,” *Trabajo Fin de Máster, EPS - UAM*, 2012.
- [6] A. C. Barboza, “Misoprostol versus Oxitocina en la inducción de la labor de parto y la maduración cervicouterina,” *Revista Médica de Costa Rica y Centroamérica*, vol. 66, no. 587, pp. 53–59, 2009.
- [7] E. Mozurkewich, J. Chilimigras, E. Koepke, K. Keeton, and V. King, “Indications for induction of labour: a best-evidence review,” *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 116, no. 5, pp. 626–636, 2009.
- [8] A. C. of Obstetricians, Gynecologists *et al.*, “Clinical Management Guidelines for Obstetrician-Gynecologists,” *ACOG Practice Bulletin. Obstetrics & Gynecology*, vol. 114 (2 Part 1), no. 107, pp. 386–397, 2009.
- [9] R. González-Boubeta and C. Cid-González, “Maduración cervical: aceleración de un proceso natural,” *Matronas Prof*, vol. 8, no. 1, pp. 24–9, 2007.

- [10] R. Altisent and M.-T. Delgado-Marroquín, “La doble faz de la técnica: informatización y confidencialidad en la historia clínica,” *Atención primaria*, vol. 34, no. 3, pp. 115–116, 2004.
- [11] J. Carnicero, “De la historia clínica a la historia de salud electrónica,” 2003.
- [12] M. Ponseti Socías, “Experiencia de las TIC en el Hospital Son Llátzer,” *Todo hospital*, no. 248, pp. 387–397, 2008.
- [13] R. L. Katz, *El papel de las TIC en el desarrollo*. Raul Katz, 2009, vol. 19.
- [14] L. básica reguladora de la Autonomía, “del paciente y de derechos y obligaciones en materia de información y documentación clínica. ley 41/2002, de 14 de noviembre,” *Boletín Oficial del Estado*, no. 274, 2002.
- [15] “Selene Clinic - Healthcare - Siemens,” 2015. [Online]. Available: http://w5.siemens.com/spain/web/es/healthcare/soluciones_it/pages/selene_clinic.aspx
- [16] *CLIPS*, NASA Johnson Space Center Std., 1984. [Online]. Available: <http://clipsrules.sourceforge.net/>
- [17] B. Buchanan and E. Shortliffe, “Rule-based expert systems: the MYCIN experiments of the Stanford heuristic programming project,” *Addison-Wesley, NY*, vol. 18, pp. 14–3, 1984.
- [18] A. J. Gonzalez and D. D. Dankel, *The engineering of knowledge-based systems: theory and practice*. Prentice-Hall, Inc., 1993.
- [19] T. R. Gruber, “A translation approach to portable ontology specifications,” *Knowledge acquisition*, vol. 5, no. 2, pp. 199–220, 1993.
- [20] S. Bechhofer, “OWL: Web ontology language,” in *Encyclopedia of Database Systems*. Springer, 2009, pp. 2008–2009.
- [21] “Protégé,” Stanford Center for Biomedical Informatics Research, 2012. [Online]. Available: <http://protege.stanford.edu/>
- [22] E. Strumbelj and I. Kononenko, “An efficient explanation of individual classifications using game theory,” *The Journal of Machine Learning Research*, vol. 11, pp. 1–18, 2010.
- [23] L. Shapley, “Contributions to the Theory of Games II,” *Value for n-person games. Annals of mathematics studies*, vol. 28, pp. 307–317, 1953.
- [24] “Hospital Universitario de Fuenlabrada,” http://www.madrid.org/cs/Satellite?language=es&pagename=HospitalFuenlabrada/Page/HFLA_home.
- [25] “Instituto de Ingeniería del Conocimiento,” <http://www.iic.uam.es/>.

- [26] “GNU Gawk,” <http://www.gnu.org/software/gawk/>.
- [27] “Ruby Programming Language,” <https://www.ruby-lang.org/>.
- [28] I. Kononenko, “Machine learning for medical diagnosis: history, state of the art and perspective,” *Artificial Intelligence in medicine*, vol. 23, no. 1, pp. 89–109, 2001.
- [29] D. A. Morales, E. Bengoetxea, P. Larrañaga, M. García, Y. Franco, M. Fresnada, and M. Merino, “Bayesian classification for the selection of in vitro human embryos using morphological and clinical data,” *Computer methods and programs in biomedicine*, vol. 90, no. 2, pp. 104–116, 2008.
- [30] L. Breiman, “Random forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [31] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA data mining software: an update,” *ACM SIGKDD explorations newsletter*, vol. 11, no. 1, pp. 10–18, 2009.
- [32] I. Kononenko, “Estimating attributes: analysis and extensions of RELIEF,” in *Machine Learning: ECML-94*. Springer, 1994, pp. 171–182.
- [33] C. Ding and H. Peng, “Minimum redundancy feature selection from microarray gene expression data,” *Journal of bioinformatics and computational biology*, vol. 3, no. 02, pp. 185–205, 2005.
- [34] A. Roy, N. Das, A. Saha, R. Sarkar, S. Basu, M. Kundu, and M. Nasipuri, “A comparative study of feature ranking methods in recognition of handwritten numerals,” in *Artificial Intelligence and Evolutionary Algorithms in Engineering Systems*. Springer, 2015, pp. 473–479.
- [35] “Python Programming Language,” <https://www.python.org/>.
- [36] R. C. Team, “A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria,” <http://www.R-project.org/>, 2013.
- [37] V. López, A. Fernández, S. García, V. Palade, and F. Herrera, “An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics,” *Information Sciences*, vol. 250, pp. 113–141, 2013.

Glosario

- *Bishop*: Escala cuantitativa desarrollada en 1964 para establecer el grado de maduración cervical (cérvix), principal factor determinante del desenlace de la inducción.
- *Borramiento*: Adelgazamiento del cuello uterino.
- *Bradycardia*: Descenso de la frecuencia cardíaca.
- *Corioamnionitis*: Inflamación o infección en la membrana que rodea al feto.
- *Cuello uterino*: Abertura del útero que se encuentra encima de la vagina.
- *Dilatación*: Ampliación de la abertura del cuello uterino.
- *Eclampsia*: Convulsiones que ocurren en el embarazo y están asociadas con niveles elevados de presión arterial.
- *Edad gestacional*: Número de semanas transcurridas desde el primer día del último periodo menstrual normal y la fecha del parto.
- *Episodio asistencial*: Intervalo de tiempo en el que se realizan actividades sanitarias por un médico para tratar un tema relativo a la salud. En este proyecto, un embarazo.
- *Feto*: Embrión que se desarrolla en el útero desde la semana 9 de embarazo hasta el final.
- *Hipertonía*: Aumento del tono uterino.
- *Líquido amniótico*: Agua en saco amniótico que rodea al feto dentro del útero de la madre.
- *Parto por cesárea*: Parto de un bebé mediante una incisión en el abdomen y útero de la madre.
- *Polisistolia*: Actividad uterina que reporta más de 5 contracciones en 10 minutos.
- *Preeclampsia*: Afección durante el embarazo en la que ocurren niveles elevados de presión arterial y proteína en la orina.
- *Taquisistolia*: Aumento de la frecuencia cardíaca.
- *Útero*: Órgano muscular ubicado en la pelvis de la mujer que contiene al feto en desarrollo y lo nutre durante el embarazo.

Acrónimos

CLIPS	C Language Integrated Production System
IR	Information Retrieval
OWL	Ontology Web Language
RDBMS	Relational Database Management System
RDF	Resource Description Framework
TIC	Tecnologías de Información y Comunicación
UUID	Universally Unique IDentifier
XML	eXtensible Markup Language



Descripción de atributos

En este anexo se describen los atributos que componen el modelo de paciente, incluyendo los histogramas de las variables relevantes (predictoras) para la clasificación y la clase con el fin de conocer la distribución de los datos. Cabe destacar que en caso de que no sea posible extraer el valor de la Historia Clínica Electrónica del paciente, se toma un valor por defecto: si es numérica, un valor nulo o, si es nominal, valor *NA*.

Atributos predictores:

- *edad_gestacional*: Edad gestacional en semanas del feto, extraída del objeto clínico más reciente.
- *edad_materna*: Edad en años, extraída del objeto clínico más reciente.
- *raza*: Raza. Los valores que puede tomar son *Caucásica*, *Afroamericana*, *Asiática*, *Latinoamericana*, *Magrebí*, *Otro*.
- *imc_inicio*: Índice de masa corporal al inicio de la gestación.
- *tabaquismo*: Indica si ha sido o es fumadora. Los valores que puede tomar son *Sí*, *No*.
- *talla_materna*: Talla en *cm*, extraída del objeto clínico más reciente.
- *tratamiento_fertilidad*: Indica si ha tomado algún tratamiento de fertilidad. Los valores que puede tomar son *Sí*, *No*.
- *partos_vaginales_previos*: Número de partos vaginales previos.

- *cesareas_previas*: Indica si ha tenido alguna cesárea previa. Los valores que puede tomar son *Sí*, *No*.
- *cervicometria*: Longitud cervical en *mm*.
- *liquido_amniotico*: Cantidad de líquido amniótico. Los valores que puede tomar son *Normal*, *Oligoamnios*, *Polihidramnios*.
- *rpm_ingreso*: Indica si ha tenido ruptura prematura de membranas al ingreso. Los valores que puede tomar son *Sí*, *No*.
- *motivo_cesarea_previa*: En caso de tener cesáreas previas, motivo de la misma. Los valores que puede tomar son *RPBF* (*riesgo pérdida bienestar fetal*), *FI* (*fracaso inducción*), *NPP* (*no progresión parto*), *DPC* (*desproporción pélvico-ceférica*), *Podálica*, *nalgas*, *Otras*.
- *complicaciones_embarazo*: Complicaciones que han surgido durante el embarazo. Los valores que puede tomar son *Diabetes gestacional con dieta*, *Diabetes gestacional insulino-dependiente*, *HTA/preeclampsia/eclampsia/EHE*, *Patología médica*, *Otras* (*fiebre, sangrado...*), *Ninguna*.
- *crecimiento_fetal*: Tipo de crecimiento del feto. Los valores que puede tomar son *Normal/PEG*, *RCIU Doppler tipo I-II*, *RCIU Doppler tipo III-IV*, *Macrosomía fetal*.
- *dbp*: Diámetro biparietal.
- *motivo_induccion*: Motivo principal por el cual un experto podría recomendar una inducción al parto. Los valores que puede tomar son *GCP*, *RPM*, *RCIU*, *PEG*, *Oligoamnios*, *Alteración bienestar fetal*, *Diabetes*, *HTA/preeclampsia/eclampsia/EHE*, *Patología médica materna*, *Otras*.
- *peso_fetal_estimado*: Peso estimado en *gr* del feto.
- *bishop_ingreso*: Índice de Bishop al ingreso.
- *numero_fetos*: Número de fetos en la gestación.

Atributos de resultado:

- *id*: Número identificador del paciente del tipo UUID (*Universally Unique IDentifier*).
- *induccion_parto*: Indica si se ha inducido el parto o no. Los valores que puede tomar son *Sí*, *No*.
- *tipo_induccion*: En caso de tener inducción, el medicamento utilizado. Los valores que puede tomar son *Oxitocina*, *Propess*, *Misofar*, *Otros*.

- *oxitocina_maxima*: Número de dosis máxima de oxitocina.
- *oxitocina_horas*: Número de horas con oxitocina.
- *propess_horas*: Número de horas con propess.
- *misofar_dosis*: Número de dosis de misofar.
- *tiempo_inicio_parto*: Minutos hasta inicio de parto.
- *tiempo_fin_parto*: Minutos desde inicio de parto a expulsivo.
- *horas_induccion*: Minutos totales de inducción.
- *alteraciones_fcf*: Indica si ha tenido alteraciones de la frecuencia cardiaca fetal. Los valores que puede tomar son *Sí*, *No*.
- *hipertonias*: Indica si ha tenido hipertonías o taquisistolias. Los valores que puede tomar son *Sí*, *No*.
- *epidural*: Indica si se le ha puesto epidural durante el parto. Los valores que puede tomar son *Sí*, *No*.
- *tipo_parto*: Tipo de parto. Los valores que puede tomar son *PE* (*parto eutócico*), *Ventosa*, *Fórceps*, *Espátulas*, *Cesárea*.
- *motivo_parto_instrumental*: Motivo de indicación de parto instrumental. Los valores que puede tomar son *RPBF* (*riesgo pérdida bienestar fetal*), *Alivio Expulsivo*, *Otras*.
- *motivo_cesarea*: Motivo de indicación de cesárea. Los valores que puede tomar son *RPBF* (*riesgo pérdida bienestar fetal*), *FI* (*fracaso inducción*), *NPP* (*no progresión parto*), *DPC* (*desproporción pélvico-cefálica*), *Podálica*, *nalgas*, *Otras*.
- *ph_cordon_arterial*: Valor del pH del cordón arterial.
- *apgar1*: Valor del test de Apgar al minuto 1 de nacer.
- *apgar5*: Valor del test de Apgar a los 5 minutos de nacer.
- *peso_neonatos*: Peso en *gr* del feto.
- *tipo_reanimacion*: Tipo de reanimación. Los valores que puede tomar son *0*, *I-II*, $\geq III$.
- *traslado_neonato*: Lugar donde trasladan al recién nacido.
- *dia_parto*: Fecha del parto.
- *hora_parto*: Hora del parto.
- *fecha_entrada*: Fecha de entrada al registro en el hospital.

Variable objetivo:

- *resultado*: Clase predicha por el clasificador. Los valores que puede tomar son *No inducción*, *Inducción*, *Cesárea*.

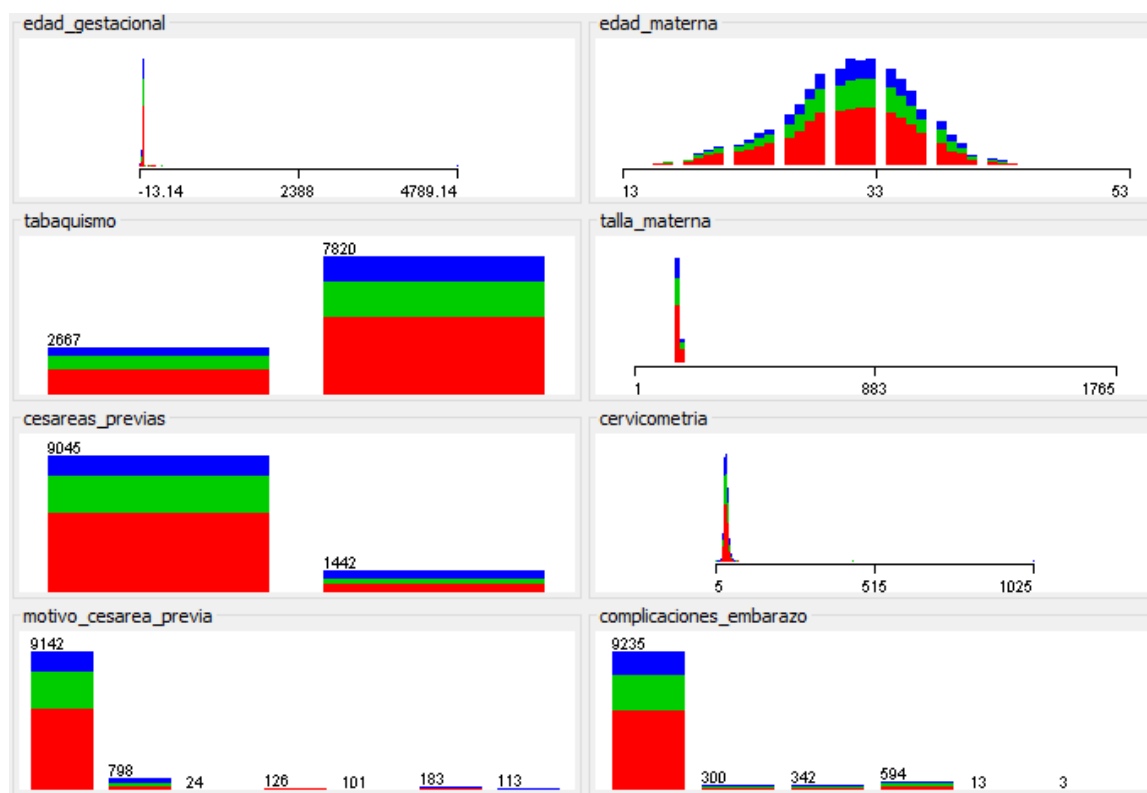


Figura A.1: Histogramas de variables predictoras del conjunto de datos INDUCESS (Parte 1). Aparece el número de instancias que toman cada valor del atributo del eje x , en color rojo la clase *No inducción*, en verde la clase *Inducción* y en azul la clase *Cesárea*.

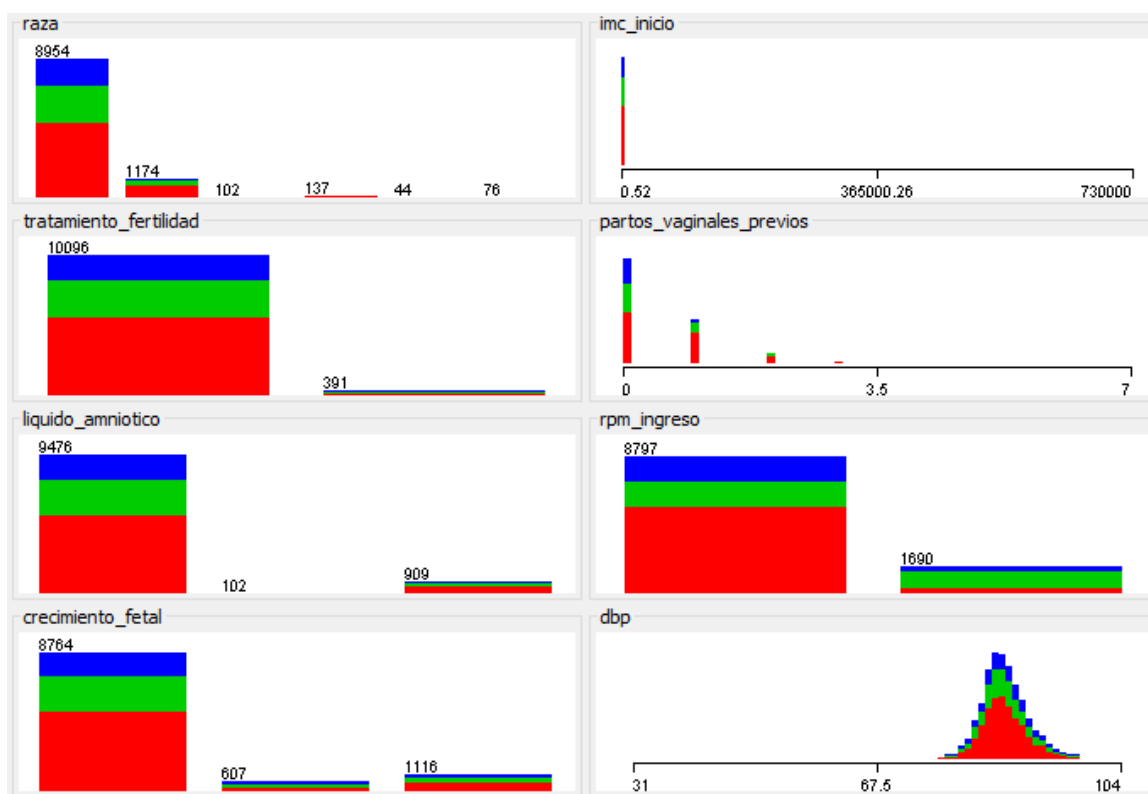


Figura A.2: Histogramas de variables predictoras del conjunto de datos INDUCESS (Parte 2). Aparece el número de instancias que toman cada valor del atributo del eje x , en color rojo la clase *No inducción*, en verde la clase *Inducción* y en azul la clase *Cesárea*.

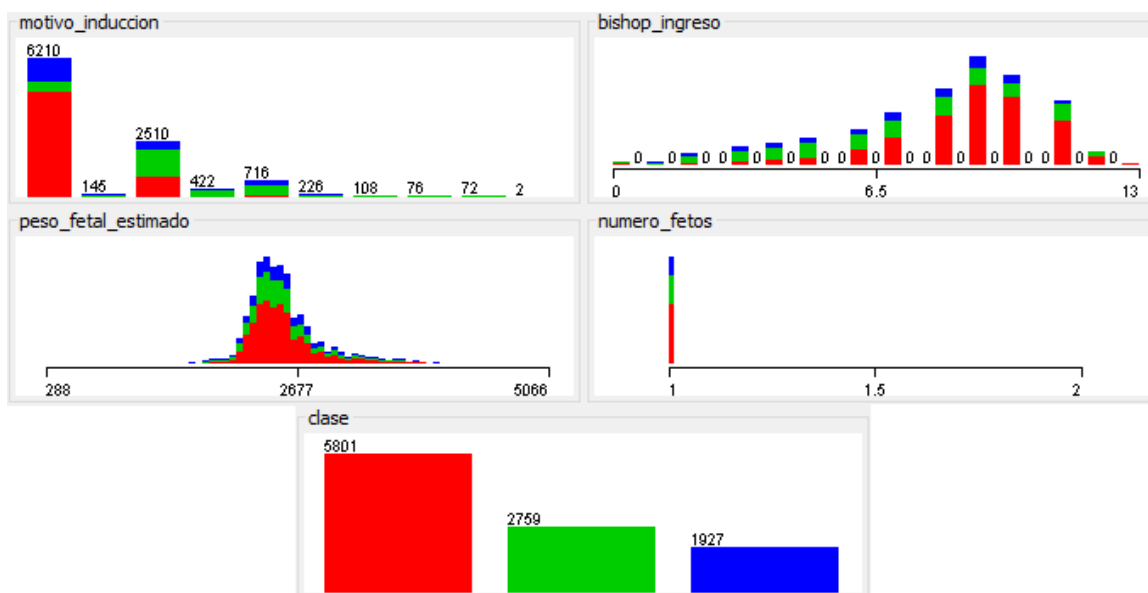


Figura A.3: Histogramas de variables predictoras y clase del conjunto de datos INDUCESS (Parte 3). Aparece el número de instancias que toman cada valor del atributo del eje x , en color rojo la clase *No inducción*, en verde la clase *Inducción* y en azul la clase *Cesárea*.

B

Pantallas página web

A continuación, se van a presentar las imágenes correspondientes a la aplicación web implementada. En la Figura B.1 se muestra la página principal completa, desde la que se puede acceder a todas las pantallas de la aplicación.

INDUCCESS Inicio Sistema de Validación Estadísticas Equipo

Bienvenid@ a Induccess

Modelo predictivo de inducción al parto

[Entrar](#)

Modelo predictivo de inducción al parto

El principal objetivo de este proyecto es diseñar un modelo predictivo del éxito de un procedimiento de inducción de trabajo de parto y construir una herramienta de ayuda a la toma de decisiones basada en:

Sistema para inferir decisiones

Se apoya en reglas de decisión provenientes de un panel de expertos en Obstetricia.

Modelos de inferencia basados en algoritmos

Se desarrollan y aplican técnicas de Big Data y aprendizaje automático.

Explicación de la clasificación

Se contribuye a hacer más transparentes los modelos considerados basándose en la teoría de juegos cooperativos.

Equipo de investigación

Induccess es una colaboración entre el Hospital Universitario de Fuenlabrada y el Entorno Salud del Instituto de Ingeniería del Conocimiento

Hospital Universitario de Fuenlabrada
Red pública de la Comunidad de Madrid

Dra. María Teulón
Servicio de Ginecología y Obstetricia

Luis Lechuga
Sistemas de Información

Carmen Yelo
Colaboradora residente

Marta Dávila
Colaboradora residente

Instituto de Ingeniería del Conocimiento
Expertos en análisis de datos y Big Data

Cristina Pruenza
Health and Energy Predictive Analytics

Julia Díaz
Directora de Innovación Health and Energy Predictive Analytics

Ana González
Health and Energy Predictive Analytics

Todo este equipo está trabajando en la realización de este proyecto que pretende mejorar la capacidad predictiva de éxito de una inducción del parto y, con ello a su vez, mejorar los resultados en salud, costes y calidad percibida del procedimiento.

Contacta con nosotros

Si quieres saber más sobre Induccess escríbenos:

Nombre y apellidos

Email

Cuéntanos...

[Enviar](#)

Figura B.1: Pantalla inicial completa del sistema de validación.

En la Figura B.2 describe muy brevemente el objetivo del sistema. Aparece un buscador desde el cual se puede introducir el identificador de un paciente concreto para revisar o, por el contrario, el sistema muestra uno aleatoriamente. Seleccionando el botón ‘Continuar’ se accede a la pantalla desde la que se puede visualizar la Historia Clínica y los atributos del paciente dado (Figura 3.4).

The screenshot shows a web application interface. At the top is a dark navigation bar with the logo 'INDUCESS' on the left and four links: 'Inicio', 'Sistema de Validación', 'Estadísticas', and 'Equipo'. The main content area has a title 'Sistema de Validación' in bold. Below the title is a paragraph: 'A partir de los datos anonimizados de la Historia Clínica Digital del Hospital Universitario de Fuenlabrada se obtiene información suficiente para poder inferir decisiones sobre la inducción del trabajo de parto.' This is followed by a smaller line of text: 'Con esta herramienta se va a validar el sistema de información construido, por favor, introduzca el identificador del paciente que desea revisar y/o pulse 'Continuar':'. Below this text is a text input field containing the alphanumeric string 'f3g2c479-eb14-3ed0-9664-2b8cce401770'. Underneath the input field are two blue buttons: 'Continuar' and 'Salir'. At the bottom of the page is a dark footer bar. On the left, it contains the text '© Instituto de Ingeniería del Conocimiento Design: TEMPLATED & IIC'. On the right, there are three circular social media icons for Twitter, Google+, and LinkedIn.

Figura B.2: Pantalla de acceso al del sistema de validación.

Como se ha explicado en el apartado correspondiente, en caso de que el paciente sea validado, la aplicación muestra el resultado propuesto por el sistema experto basado en reglas.



Figura B.3: Pantalla de resultado de INDUCESS.

En la Figura B.4 se muestran datos relativos al conjunto de datos: número de pacientes, número de informes, número de notas no estructuradas, número de formularios y número de peticiones de pruebas registradas en la base de datos; y datos del proceso de validación: pacientes validados, rechazados y pendientes de revisar. Estos últimos datos además se muestran mediante un gráfico circular animado, de manera que con un vistazo se pueda saber la proporción de datos revisados.



Figura B.4: Pantalla de estadísticas del sistema de validación.

Finalmente, en la Figura B.5 se describe el proyecto y el equipo de trabajo que compone este proyecto.

INDUCESS
Inicio
Sistema de Validación
Estadísticas
Equipo

Equipo de trabajo

Este proyecto surge de la necesidad clínica de resolver una situación de incertidumbre que se presenta con frecuencia en la práctica asistencial. La **inducción del trabajo de parto** es un procedimiento extraordinariamente común empleado en alrededor de un 20% de las gestaciones. A pesar de que la inducción sea una de las intervenciones mejor estudiadas dentro del campo de la Obstetricia y existan numerosos estudios publicados acerca de ella, todavía no se han identificado buenos **factores predictivos** del éxito de este procedimiento.

Este hecho provoca que hasta el momento no existan herramientas que apoyen al médico experto en su criterio a la hora de tomar la decisión de realizar una inducción, viéndose obligado a decidir en base a ciertas características de la madre y el feto, pero que en el caso de ser una decisión ilógica e inaceptable podría ponerles en riesgo. Bajo esta situación, resulta interesante conocer a priori la posibilidad de **éxito de la inducción** para desestimar aquellas inducciones con alta probabilidad de fracaso, evitando así procedimientos innecesarios o posponiendo la finalización si fuese posible.

El principal objetivo de este proyecto es diseñar **un modelo predictivo del éxito** de un procedimiento de inducción del trabajo de parto y construir una herramienta de ayuda a la toma de decisiones basada en dos enfoques. Por un lado, un sistema para inferir decisiones que se apoya en **reglas de decisión** provenientes de un panel de expertos en Obstetricia y, por otro, modelos de inferencia basados en algoritmos y **técnicas de aprendizaje automático**, a la vez que se da continuidad a una vía de investigación consistente en la posible explicación de la clasificación obtenida.

Con la realización de este proyecto se evita llevar a cabo aquellos procedimientos con alta probabilidad de fracaso, mejorando la capacidad predictiva de éxito de una inducción del parto y, en general, los resultados en salud, costes y calidad percibida del procedimiento.

Hospital Universitario de Fuenlabrada

El Hospital Universitario de Fuenlabrada está integrado en la red pública de la Comunidad de Madrid y ofrece cobertura a una población superior a 213.000 personas correspondientes a los municipios de Fuenlabrada, Moraleja de Enmedio y Humanes de Madrid.

Para prestar a todos los usuarios la mejor asistencia posible, siguiendo el modelo de calidad europeo EFQM, el Hospital Universitario de Fuenlabrada ha puesto en marcha un Plan de Mejora, incardinado dentro de su Plan Estratégico, que se ha centrado en cinco grandes líneas: **la adecuación tecnológica** y de la oferta de servicios a las necesidades de la población, el trato adecuado y respetuoso a los pacientes, la política de excelencia hacia las personas como expresión del compromiso de la Institución con sus profesionales, el logro de una asistencia segura, ágil, resolutive, y eficiente y **la innovación y el desarrollo en sistemas de información** para la mejora del funcionamiento del Centro en todos sus ámbitos de actuación.

Como fruto de todo el trabajo el hospital ha obtenido, entre otros, el galardón Hospital Amigo de los Niños (UNICEF) para el desarrollo y fomento de la lactancia materna y el Most Wired International Citation of Merit, como organización excelente en la gestión de las tecnologías de información en Sanidad.

Instituto de Ingeniería del Conocimiento - IIC

El Instituto de Ingeniería del Conocimiento (IIC) es una entidad privada de I+D+i especializada en la extracción de conocimiento a partir de **grandes volúmenes de datos** heterogéneos (Big Data) y la optimización de procesos empresariales en campos como la banca, la salud, los medios sociales, la energía y la gestión del talento en RR. HH.

Se autofinancia a través de la venta de los productos y servicios resultantes de sus propias investigaciones, demostrando así que investigación y comercialización son compatibles y rentables, y reinvertiendo los beneficios en más innovación e investigación. Además, el IIC apuesta por el desarrollo de la **propiedad intelectual** como fórmula para incrementar la productividad y la competitividad en los mercados globales.

El equipo del IIC, siempre a disposición de sus clientes, lo integran profesionales **altamente cualificados**: psicólogos y psicómetras, lingüistas, matemáticos, físicos, informáticos e ingenieros expertos en sus disciplinas y especializados en la creación de modelos y tecnologías de análisis de datos.

© Instituto de Ingeniería del Conocimiento. Design: TEMPLATED & IIC


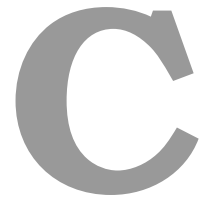



Figura B.5: Pantalla de descripción del equipo del proyecto.



Reglas de decisión

Las reglas de decisión que conforman la base de conocimiento del sistema experto se detallan a continuación.

1. **IF** *motivo_induccion*! = *NA* **and** *cesareas_previas* == *Sí*
THEN *Cesárea*
2. **IF** *crecimiento_fetal* == *RCIU Doppler tipo III-IV*
THEN *Cesárea*
3. **IF** *crecimiento_fetal* == *RCIU Doppler tipo I-II*
THEN *Inducción*
4. **IF** *motivo_induccion*! = *NA* **and** *crecimiento_fetal* == *Normal/PEG*
THEN *Inducción*
5. **IF** (*motivo_induccion* == *RPM* **or** *rpm_ingreso* == *Sí*) **and** *edad_gestacional* > 34 **and** (*numero_fetos* == 1 **or** *numero_fetos* == 2)
THEN *Inducción*
6. **IF** (*motivo_induccion* == *RPM* **or** *rpm_ingreso* == *Sí*) **and** *edad_gestacional* < 34 **and** *complicaciones_embarazo* == *Otras (fiebre, sangrado...)*
THEN *Inducción*

7. **IF** (*motivo_induccion*== *RPM* **or** *rpm_ingreso*==*Sí*) **and** *edad_gestacional*< 34 **and** *complicaciones_embarazo*==*Ninguna*
THEN *No inducción*
8. **IF** *oxitocina_horas*≥ 16 **and** *misofar_dosis*≥ 4 **and** *bishop_ingreso*< 6 **and** (*motivo_cesarea*==*FI* (*fracaso inducción*) **or** *motivo_cesarea*==*NPP* (*no progresión parto*))
THEN *Cesárea*
9. **IF** *bishop_ingreso*> 6 **and** *tipo_parto*==*PE* (*parto eutócico*)
THEN *No inducción*
10. **IF** *bishop_ingreso*> 6 **and** *tipo_induccion*==*Oxitocina*
THEN *Inducción*
11. **IF** *bishop_ingreso*< 6 **and** (*edad_gestacional*< 34 **or** *complicaciones_embarazo*== *EHE* **or** *motivo_induccion*== *EHE* **or** *motivo_induccion*== *RCIU* **or** *cesareas_previas*==*Sí* **or** *num_fetos*== 2) **and** *tipo_induccion*==*Propess*
THEN *Inducción*
12. **IF** *bishop_ingreso*< 6 **and** *edad_gestacional*> 34 **and** **not**(*complicaciones_embarazo*== *EHE* **or** *motivo_induccion*== *EHE* **or** *motivo_induccion*== *RCIU* **or** *cesareas_previas*==*Sí* **or** *num_fetos*== 2) **and** *tipo_induccion*==*Misofar*
THEN *Inducción*
13. **IF** (*complicaciones_embarazo*==*Diabetes gestacional con dieta* **or** *complicaciones_embarazo*==*Diabetes gestacional insulino dependiente*) **and** *crecimiento_fetal*==*Macrosomía fetal* **and** *peso_fetal_estimado*> 4000
THEN *Cesárea*
14. **IF** (*complicaciones_embarazo*==*Diabetes gestacional con dieta* **or** *complicaciones_embarazo*==*Diabetes gestacional insulino dependiente*) **and** *peso_fetal_estimado*> 4500
THEN *Cesárea*
15. **IF** (*complicaciones_embarazo*==*Diabetes gestacional con dieta* **or** *complicaciones_embarazo*==*Diabetes gestacional insulino dependiente*) **and** *peso_fetal_estimado*< 4500
THEN *Inducción*

16. **IF not**(*complicaciones_embarazo*==*Diabetes gestacional con dieta* **or** *complicaciones_embarazo*==*Diabetes gestacional insulino dependiente*) **and** *peso_fetal_estimado*> 5000
THEN *Cesárea*
17. **IF not**(*complicaciones_embarazo*==*Diabetes gestacional con dieta* **or** *complicaciones_embarazo*==*Diabetes gestacional insulino dependiente*) **and** *peso_fetal_estimado*< 5000
THEN *Inducción*
18. **IF** (*motivo_induccion*== *GCP* **or** *edad_gestacional*> 41) **and** *numero_fetos*== 1 **and** *complicaciones_embarazo*== *Ninguna*
THEN *Inducción*
19. **IF** (*motivo_induccion*== *GCP* **or** *edad_gestacional*> 41) **and** *numero_fetos*== 2 **and** *cesareas_previas*==*Sí*
THEN *Cesárea*
20. **IF** (*motivo_induccion*== *GCP* **or** *edad_gestacional*> 41) **and** *numero_fetos*== 2 **and** *cesareas_previas*==*No*
THEN *Inducción*
21. **IF** (*motivo_induccion*== *GCP* **or** *rpm_ingreso*==*Sí*) **and** *numero_fetos*== 2 **and** *cesareas_previas*==*Sí* **and** *bishop_ingreso*< 6
THEN *Cesárea*
22. **IF** (*motivo_induccion*== *GCP* **or** *rpm_ingreso*==*Sí*) **and** *numero_fetos*== 2 **and** *cesareas_previas*==*No*
THEN *Inducción*
23. **IF** *bishop_ingreso*≤ 6 **and** *partos_vaginales_previos*== 0 **and** *talla_materna*< 155
THEN *Cesárea*
24. **IF** *bishop_ingreso*≤ 2 **and** *partos_vaginales_previos*≥ 1 **and** (*talla*≤ 161 **or** *peso*> 3500)
THEN *Cesárea*
25. **IF** *bishop_ingreso*≤ 6 **and** *cesareas_previas*==*Sí* **and** (*talla*≤ 161 **or** *peso*> 3500)
THEN *Cesárea*
26. **IF** *bishop_ingreso*≤ 6
THEN *Cesárea*

27. **IF** *bishop_ingreso* > 6
 THEN *Inducción*
28. **IF** *peso_fetal_estimado* > 4000
 THEN *Cesárea*
29. **IF** *imc_inicio* ≤ 18
 THEN *Cesárea*
30. **IF** *imc_inicio* ≥ 36
 THEN *Cesárea*
31. **IF** *talla_materna* ≤ 160
 THEN *Cesárea*
32. **IF** *edad_gestacional* < 34
 THEN *Cesárea*
33. **IF** *cesareas_previas* == *Sí* **and** (*motivo_cesarea_previa* == *FI* (*fracaso inducción*) **or** *motivo_cesarea_previa* == *NPP* (*no progresión parto*))
 THEN *Cesárea*
34. **IF** *raza* == *Afroamericana*
 THEN *Cesárea*
35. **IF** *motivo_induccion* != *NA* **and** *motivo_cesarea* == *Otras*
 THEN *Cesárea*



Matrices de confusión

A lo largo de este anexo se presentan las matrices de confusión resultantes de cada ejecución de los algoritmos utilizados en la Sección 6.

Tabla D.1: Matrices de confusión del Modelo 1 (Falsos predictores).

	Real/Predicha	No inducción	Inducción	Cesárea
Naive Bayes	No inducción	5.600	82	119
	Inducción	13	2.724	22
	Cesárea	18	15	1.894
J48	No inducción	5.801	0	0
	Inducción	0	2.759	0
	Cesárea	0	0	1.927
Multilayer Perceptron	No inducción	5.801	0	0
	Inducción	0	2.759	0
	Cesárea	0	0	1.927
SVM	No inducción	5.603	77	121
	Inducción	14	2.710	35
	Cesárea	34	40	1.853
Random Forest	No inducción	5.620	26	155
	Inducción	219	2.459	81
	Cesárea	267	79	1.581

Tabla D.2: Matrices de confusión del Modelo 2 (Atributos predictores).

	Real/Predicha	No inducción	Inducción	Cesárea
Naive Bayes	No inducción	3.932	933	936
	Inducción	283	2.050	426
	Cesárea	490	598	839
J48	No inducción	5.192	385	224
	Inducción	552	1.899	308
	Cesárea	934	532	461
Multilayer Perceptron	No inducción	4.975	404	422
	Inducción	499	1.912	348
	Cesárea	609	539	779
SVM	No inducción	5.048	566	187
	Inducción	725	1.861	173
	Cesárea	1.014	604	309
Random Forest	No inducción	5.080	377	344
	Inducción	497	1.973	289
	Cesárea	818	520	589

Tabla D.3: Matrices de confusión del Modelo 3 - Atributos predictores tras ReliefF.

	Real/Predicha	No inducción	Inducción	Cesárea
Naive Bayes	No inducción	5.092	569	140
	Inducción	816	1.761	182
	Cesárea	1.036	677	214
J48	No inducción	4.991	760	50
	Inducción	646	1.981	132
	Cesárea	1.048	747	132
Multilayer Perceptron	No inducción	4.942	644	215
	Inducción	844	1.666	249
	Cesárea	630	592	705
SVM	No inducción	4.787	1.012	2
	Inducción	473	2.272	14
	Cesárea	1.048	856	23
Random Forest	No inducción	4.852	662	287
	Inducción	826	1.506	427
	Cesárea	1.032	556	339

Tabla D.4: Matrices de confusión del Modelo 3 - Atributos predictores tras mRMR.

	Real/Predicha	No inducción	Inducción	Cesárea
Naive Bayes	No inducción	5.228	480	93
	Inducción	1.024	1.636	99
	Cesárea	1.115	634	178
J48	No inducción	4.798	807	196
	Inducción	581	2.028	150
	Cesárea	841	705	381
Multilayer Perceptron	No inducción	4.843	696	262
	Inducción	728	1.765	266
	Cesárea	556	593	778
SVM	No inducción	4.634	991	176
	Inducción	456	2.169	134
	Cesárea	831	723	373
Random Forest	No inducción	4.784	770	247
	Inducción	678	1.851	230
	Cesárea	855	625	447

Tabla D.5: Matrices de confusión del Modelo 3 - Atributos predictores tras GainRatioAttributeEval.

	Real/Predicha	No inducción	Inducción	Cesárea
Naive Bayes	No inducción	5.056	413	332
	Inducción	512	2.087	160
	Cesárea	848	623	456
J48	No inducción	5.168	439	194
	Inducción	507	2.189	63
	Cesárea	940	732	255
Multilayer Perceptron	No inducción	5.188	409	204
	Inducción	509	2.070	180
	Cesárea	966	616	345
SVM	No inducción	5.261	439	101
	Inducción	516	2.218	25
	Cesárea	1.082	751	94
Random Forest	No inducción	5.142	423	236
	Inducción	510	2.151	98
	Cesárea	902	664	361

Tabla D.6: Matrices de confusión del Modelo 3 - Atributos predictores tras InfoGainAttributeEval.

	Real/Predicha	No inducción	Inducción	Cesárea
Naive Bayes	No inducción	4.083	949	769
	Inducción	273	2.098	388
	Cesárea	413	661	853
J48	No inducción	5.308	358	135
	Inducción	578	2.003	178
	Cesárea	975	591	361
Multilayer Perceptron	No inducción	5.180	370	251
	Inducción	561	1.896	302
	Cesárea	617	541	769
SVM	No inducción	5.227	431	143
	Inducción	519	2.117	123
	Cesárea	905	617	405
Random Forest	No inducción	4.913	416	472
	Inducción	557	1.780	422
	Cesárea	865	482	580

Tabla D.7: Matrices de confusión del Modelo 3 - Atributos predictores tras CorrelationAttributeEval.

	Real/Predicha	No inducción	Inducción	Cesárea
Naive Bayes	No inducción	4.810	515	476
	Inducción	406	2.016	337
	Cesárea	766	551	610
J48	No inducción	5.317	400	84
	Inducción	547	2.077	135
	Cesárea	1.045	660	222
Multilayer Perceptron	No inducción	5.181	357	263
	Inducción	574	1.890	295
	Cesárea	631	518	778
SVM	No inducción	5.197	436	168
	Inducción	505	2.143	111
	Cesárea	916	653	358
Random Forest	No inducción	5.201	382	218
	Inducción	575	1.988	196
	Cesárea	937	606	384